257
1.9.70

VOLUME 1. NO. 1 NOVEMBER, 1958

EDUCATIONAL RESEARCH

A Review for Teachers and all concerned with Progress in Education

BUREAU OF EDNL. & PSYL. RESEARCH
LIBRARY

Published for
THE NATIONAL FOUNDATION FOR EDUCATIONAL
RESEARCH IN ENGLAND AND WALES

by

NEWNES EDUCATIONAL PUBLISHING CO. LTD.
Tower House, Southampton Street, Strand, London, W.C.2

EDITORIAL

EDITORIALS will not be a frequent or regularly recurring feature of this review but the first issue is evidently a special occasion. It seems appropriate to say something of the intentions in the minds of those responsible, something of the purposes we hope will be served and something of future editorial policy.

Many people in the education service are aware of the increasing gap between the worlds of research and the schools. With the best will in the world it is difficult for teachers to keep abreast of the new knowledge which is daily acquired by research workers in universities and in organizations like our own both in this country and abroad. The sheer volume of published work is more than one person can read and digest. Moreover, refinements in statistical and experimental methods, essential factors in the soundness of the work done, inevitably increase the difficulty of understanding for those who are not themselves professional research workers.

It is our purpose to bridge this gap. The articles in each number of the review will be substantial summaries of work in various research fields. We have no intention of 'popularizing' in the sense that articles are written down, or that they avoid such technical terminology as is necessary. The papers are intended to have the nature of comprehensive and authoritative statements of the principal facts and conclusions derived from research and are intended to provide our professional colleagues with the basis on which they can take practical decisions.

Certain topics are of such broad importance that there is much information available —more than can reasonably be digested into one paper. Such topics will, therefore, be treated in serial fashion, under a general heading such as *The Teaching of Reading* and the aim is that over a number of issues of the review a thorough study will be made. Such series of articles are planned to deal with practical problems of teaching method and organization. Single articles are in general devoted to more general problems—such as that by Professor Vernon in this issue.

In addition to the articles themselves, the review will have two other continuing features. The first of these is the abstract section in which from time to time major pieces of research work or research papers from foreign journals which may not be readily accessible in this country, will be fully summarized. The second is the bibliographical material which will either appear as an appendix to each article or in the form of a select annotated bibliography of the kind published in this issue.

We believe that a review of this kind is urgently needed by all those concerned with education; and we think that the proposed style and policy are approximately right. However, *Educational Research* is intended as a tool for those who read it and we shall be glad to have criticisms, suggestions for articles or indeed any comment which will enable us to give a better service.

W.D.W.

[illegible] Research
[illegible] THE
Dated 1. 9. 76
Acc. No. J 257

A NEW LOOK AT INTELLIGENCE TESTING

PHILIP E. VERNON

(Professor of Educational Psychology, University of London, Institute of Education).

Introduction.

TEACHERS who have followed the controversies over the past few years in the National Foundation's *Bulletin*, or in the periodical *Education*, and the *Times Educational Supplement*, may be pardoned if they have reached the conclusion that psychologists don't know their own minds. Apparently many of the fundamental principles of intelligence and intelligence testing which they imbibed from Ballard, Nunn, Hughes or Schonell are nowadays suspect, although still stoutly defended by a few authorities such as Professor Burt.

It might almost seem as though the whole raison d'être of using intelligence tests in educational selection and guidance had been demolished. That conclusion—I hope to show—is certainly not justified, though I would agree that many facts have emerged from research over the past 30 years or so which do indicate the need for some re-orientation in our theories of intelligence and in our practical application of tests. I should also claim that this ' new look ' in intelligence testing represents a change of emphasis rather than a revolution ; for actually much of what follows is in accord with what Burt, Schonell and I myself were teaching in the 1920's and 30's.

At the same time I would admit that psychologists are somewhat to blame for allowing a rather over-simplified and over-optimistic view of intelligence testing to become absorbed by student-teachers and by educationists generally—that is by people who could hardly be expected to keep in mind all the statistical snags and technical qualifications of which psychologists themselves should be aware. What are the essentials of this view which can no longer be accepted? They can be summarised as follows :—

The I.Q. as measured by standard intelligence tests, applied in a standard manner, measures the child's innate intellectual capacity which is determined by his genes. This general ability or *g* is the main factor underlying achievement in any direction, particularly in the educational sphere. Moreover, since this index of intelligence is very little affected by upbringing or environment, it enables us to predict, even in the junior school, the child's present and ultimate capacity for acquiring school learning and his suitability for a career involving high (or low) intelligence. If he is backward educationally and his I.Q. is low, we need seek no further explanation. But if his I.Q. is superior to his attainment, there are probably environmental or temperamental causes, or defects in schooling, which should be remedied so that he can work ' up to capacity '.

Variations In, and Diverse Kinds of, Intelligence.

The first aspect of this view which I must contest is the implicit assumption that intelligence is a single, clear-cut mental power or faculty, or in other words that any one child will get the same I.Q. whatever test he is submitted to. In fact it is only partly true that different intelligence tests measure the same thing. They do overlap or correlate fairly highly with one another, but somewhat different results will certainly be obtained from, say, Moray House and National Foundation tests, more different from a verbal group test and individual Terman-Merrill, and still more different from verbal tests and non-verbal ones based on pictures, diagrams, or practical materials. Unfortunately it is far too common for teachers to apply one test and to enter the results under I.Q. on the record card, as though this was fixed immutably, regardless of the test used or the age of testing.

Now most children, when tested several times, do tend to score in the same band or region of I.Q. The average change when the tests are similar in character should be about 6 points, though this may rise to 10 points if the tests are more dissimilar, or if the period of time amounts to several years. Some are still more stable, but others are much more variable, and occasional gains or losses of 30 or even 40 points do occur. Suppose we regard 15 points as a big change, then about 1 in 9 will alter this much or more over a month or so with similar tests, and 1 in 6 will do so over several years or with dissimilar tests*.

A good deal of this variability is fortuitous : one test may have been given under noisy or otherwise unsuitable conditions ; the tester may add up the marks wrongly ; the children may have had some coaching ; the test norms or standards may be inaccurate ; or the spread of I.Q's—as measured by the standard deviation—may be abnormal ; any of these may differ between one test and another, and therefore upset the results to some extent. But even allowing for these sources of unreliability, it is clear that tests composed of different materials—verbal, numerical, spatial, etc.—or tests involving different functions such as inductive or deductive reasoning, memorising, fluency, etc. or using different sorts of items—multiple-choice vs. creative-response, speeded vs. unspeeded—all these measure somewhat different abilities.

Spearman was entirely justified in his claim that a common underlying factor *g* runs through all our mental abilities, and that this *g* is most strongly present in the higher thinking processes such as comprehension, reasoning, abstracting and generalising, or grasping relations. But he did not, like Burt and like the American factorists such as Thurstone, sufficiently realise the diversity as well as the unitariness of abilities.

This means then that intelligence is best regarded as a fluid collection of infinitely varied thinking abilities, and *g* is simply an average of whatever abilities the tester likes to include ; it cannot be pinned down to any specified intellectual function. Thus different tests or test batteries will inevitably yield somewhat different results, although they have much in common. There is no means of deciding that one test which samples one particular set of abilities is a better test of a hypothetical entity—intelligence—than is another test which samples a somewhat different set of abilities.

A further implication that many psychologists have drawn is that it would be preferable to give up the term intelligence, and instead build up batteries of tests which will best predict the particular kinds of intelligence in which we happen to be interested.

---

*These figures assume correlations of 0.90 and 0.70 respectively, also that the tests compared are accurately standardised with the same standard deviation of 15. If the norms differ, large variations may be more frequent.

For example we would have a set of Scholastic Aptitude tests for predicting educational success in predominantly linguistic courses, and a partly overlapping, partly different, set for predicting technical aptitude, and yet another for assessing brightness and quickness-in-the-uptake in everyday life, and so on. It is noticeable that the National Foundation issues tests called Verbal or Non-Verbal Tests rather than intelligence tests ; and Moray House has likewise recently transformed its series into Verbal Reasoning Tests.

In the scholastic field, there is no need to confine ourselves to the rather stereotyped and narrow range of types of item conventionally included in intelligence tests. Better predictions might be obtained from tests aimed at different kinds of reasoning, flexibility of mind, critical judgment, original thinking, and other partially distinguishable intellectual faculties.

Environmental Effects On Intelligence.

Let us now turn to changes in children's intelligence as they grow older. Such changes, attributable to environmental influences or to developments in the children's personalities, certainly do occur ; though they are probably less frequent and less marked than would appear merely from observing successive sets of I.Q's. because discrepancies so often arise—as we have seen—either from variations in the factorial content of the tests or from various fortuitous circumstances. However there is sufficient evidence of the effects of social class, upbringing and education to negate the view that intelligence is wholly inborn and therefore unalterable.

First, though, we should distinguish between improvements due to environmental factors and I.Q. gains brought about by coaching or practice on intelligence tests. These latter gains are certainly larger than we used to suppose : they range from about 3 to 12 points depending on the type of training given and the degree of previous sophistication of the children. But their most noteworthy features are that they are brought about very rapidly by 3 hours or less of training (thereafter there is hardly any further consistent improvement), and that they are highly specific to the particular types of items and the particular form of testing. There is no general spread to other kinds of intellectual performance. The children would seem merely to have acquired a familiarity with the mechanics of the tests on which they have been coached or practised. Thus, far from regarding this improvement as training their intelligence, I would suggest that most of the coaching that goes on for ' the 11 plus " is antithetical to true growth of their all-round thinking capacities.

Extensive investigations have been carried out into the effects of good and poor foster homes on the intelligence of orphan children ; also, identical twins, who have precisely the same heredity but who have been separated and brought up in different environments, have been studied. These researches rather uniformly indicate that the difference between a good and a poor home may raise or lower the I.Q. by about 10 points, or in more extreme cases by 20 points. At the same time there is no justification for those left-wing views which try to ascribe all intelligence differences to social class and upbringing, since some studies show small but significant correlations remaining between the intelligence of orphans and that of their true parents who have *not* brought them up. Moreover it would be practically impossible to explain why children in the same family often show quite a wide range of I.Q., why professors sometimes have dull children and manual labourers bright ones, nor why orphans in a highly standardised institution do not all tend to approximate to the same dead level of intelligence, if environment were supreme. Whereas on genetic theory we would expect each brother and sister, or each orphan, to inherit a somewhat different proportion of the genes that make for good intelligence.

From twin data Burt and others have calculated that something like four-fifths of the differences among primary school children in intelligence should be attributed to hereditary causes. I do not regard that estimate as unreasonable when we consider the remarkable degree of uniformity of environment to which children are exposed up to 11 years. Certainly homes vary considerably in material circumstances and cultural stimulus. But so long as children are building up the simpler thinking capacities it is likely that most home and leisure environments provide the necessary exercise for their maturing brain processes. Indeed from 5 to 11 they are subjected to a highly standardised intellectual training at school.

But at the age when more complex and abstract thinking is developing, environments begin to differ more widely : they go to schools which exert very different degrees of intellectual pressure. By 15, half or more cease any formal education, enter largely routine jobs and indulge in leisure pursuits which make little call on intelligence ; whereas a privileged minority continue to exercise their minds to the utmost at school or university, are more likely to enter occupations that exercise their minds and to carry on cultural pursuits that make them think.

Hence I was able to show, by testing all the boys in a large city at 14 years that, after allowing for their different levels at 11+, those in the best secondary schools had gained 12 points over those in the weakest schools. The average modern school pupil had lagged 7 points behind the average grammar or technical school pupil.* Naturally, much of this difference might be due to the homes the different kinds of pupils came from rather than to the goodness of schooling as such; but it was, none the less, an environmental effect.

Similarly, Lorge in America and Husén in Sweden have tested the same individuals as children and as adults, and have shown that the later scores are definitely affected by the amount of schooling these individuals have received in the intervening period. Boys of the same I.Q. at the start, who had full secondary and further education, end up some 12 points ahead of those who left school at 14 to 15 and got no more education. Yet another piece of confirmatory evidence came from my analysis of the Progressive Matrices test scores of naval recruits of various ages during the war. This showed that the average capacity of adults to reason at this kind of test tends to decline with age; however, it stays up longer among men coming from skilled civilian occupations than among those whose civilian jobs appear to provide less intellectual stimulus.

---

*Yates and Pidgeon state that they are unable to confirm this finding. But they did not allow for the normal so-called regression effect. Had the grammar and modern schools in their follow-up had no effects on pupils' intelligence, the difference between their groups should have been 19.63 I.Q. points, if we guess the regression coefficient at 0.80. Since the obtained difference was 26.53 points, this indicates an alteration ,attributable to environment, of 7.1 points.

Such findings have tremendously important educational implications, since they show that full realisation of intellectual powers during the 'teens is often inhibited by early cessation of education, or by providing adolescents with a routine sort of schooling and conditions of work and leisure which deaden any desire to develop their minds further. While I personally favour education at the independent or maintained grammar school as doing most to draw out the intellectual faculties of the bright adolescent, I also feel a great deal of sympathy with the comprehensive school as providing more opportunities for the development of the pupil whose potentialities have been missed at 11+. The recent reports by the National Foundation and the British Psychological Society both show that the accuracy of selection at the end of primary schooling is remarkably high; in other words, that intelligence and attainment levels do remain fairly stable in the great majority. But they also agree that changes are to be expected in the ensuing years, and that therefore any rigid system of streaming or selection is bound to do injustice to at least 5 per cent. of children.

Much the same situation holds at the bottom end of the scale. A few per cent. of pupils drop so far behindhand in the primary school that it is to everyone's advantage to transfer them to special classes or to E.S.N. schools. But, partly because of the dogma of the fixity of the I.Q., it has been assumed that such pupils are incurably defective. They have become stuck in special schools and institutions. Indeed, many imbeciles have vegetated for the whole of their lives in mental hospitals, because we have failed to recognise that considerable improvements can occur in some—though certainly not in all—cases.

Recently, British psychologists, such as O'Connor, Clarke and Mundy, have been able to demonstrate significant I.Q. rises among feeble-minded patients, whose manual and other special abilities are utilised by training for appropriate jobs and who, by becoming responsible, largely self-supporting citizens escape from the deadening effects of the institutional environment. There is an important lesson here for the organisation of E.S.N. education. It is probable that a less rigid system of special classes might enable a large proportion to return to normal schooling who are at present left behind in segregated schools.

At the same time we must try to retain a balanced judgment. Not only are environmental effects on intellectual functioning rather limited in extent, but also they are by no means easy to control. We have no recipe for raising intelligence at will. We merely know, rather vaguely, that a lot of schooling is better than a little, that some schools are probably more effective than others, that cultured middle-class homes usually provide somewhat fuller realisation of potentialities than overcrowded slum homes. However, it should be possible through research to find out just what kinds of schooling, and what factors in the home, do most to stimulate all-round intellectual development.

The need for caution is also indicated by the negative results of so many experiments on transfer of training, which are quite rightly harped on in educational psychology textbooks. The ' new look ' that I am advocating does not justify a return to the old belief in formal discipline, and the pre-eminence of a classical education. One other pointer that does emerge from investigations of pre-school children is the vital part played in intellectual development by emotional factors: there is strong evidence that the young child who is emotionally secure in a democratic family environment is better able to exploit and develop his intellectual capacities than the child who is either over-protected or who feels rejected. And it may turn out that some subtle qualities of the classroom climate and the teacher's personality contribute as much, or more, to the intelligence of primary and secondary pupils as do the particular topics taught or the techniques of teaching.

The Psychology of Intellectual Development

There is another strong reason for revising our views on intelligence, namely, the intensive studies of children's psychological development that have been carried out by Gesell and his followers in America, and particularly by Piaget in Switzerland. Piaget shows that none of our mental capacities is innately given, or develops in isolation: even the baby's recognition of his rattle as a distinct object which can be moved by him or others and which makes noises, but nevertheless retains its identity, has to be built up through continuous exploration of, and experiment with, his environment. These early sensory-motor and perceptual associations form the essential basis for his later images and ideas, which can be symbolised by means of words.

Similarly, his manipulations of ideas, which we call thinking or reasoning, again depend on practice and experiment, and they are shaped and corrected by adult teaching and other environmental influences. His whole mental development is a cumulative process of interaction between his maturing potentialities and environmental opportunities and pressures. We can readily see, therefore, that in an abnormal or deprived environment, perceptual and conceptual development may be hindered, and consequently his intelligence level reduced; whereas in a home where rich experience of objects, words and logical thought are provided, and where emotional adjustment is harmonious, the growth of intelligence may be fostered.

Further valuable suggestions as to the brain mechanisms underlying mental development have been made by Hebb, the Canadian neurological psychologist. And he draws a most useful distinction between two uses of the term intelligence. On the one hand there is Intelligence A, that is, the basic potentiality for building up associations and for mental development generally—some property of the brain and central nervous system which is ultimately determined by the genes. Whereas the intelligence we see in everyday life or at school, and which our tests sample fairly effectively, is Intelligence B; and this represents the actual complexity and efficiency of mental processes reached through the cumulative process I have been describing. In other words, Intelligence A is the innate capacity to acquire Intelligence B. But we can never observe Intelligence A directly, nor measure it; it is merely a useful hypothesis. Thus, it would probably be wiser to forget about it and to concentrate on 'Intelligence-B-as-developed-by-environment.' Although, as shown above, the latter is very variegated, yet it can be reliably measured at any one time, and does stay fairly stable among a majority of normal children.

But the reason why it enables us to predict educability and future progress over the next few years is not so much because it is innately determined as because of its essentially cumulative nature. By the time children reach, say, 11 years, the interaction of their innate potentialities with environmental stimulation has brought them to a certain level of intellectual functioning, and this level will largely determine their future growth, though it may always speed up, or be inhibited, in a minority depending on the kind of intellectual and emotional experiences they meet.

Intelligence and Attainment

The question will naturally be asked—how then does intelligence differ from attainments, and the answer is—much less than we commonly suppose. For innate and external factors play very much the same role in the child's acquisition of reading or arithmetic as in the building up of his intelligence. It is very natural that teachers should tend to over-estimate the contribution that schooling makes to attainment. Yet there is much experimental evidence indicating that wide differences in the school time devoted to a subject, and differences in the methods of teaching, have relatively small effects on the amount absorbed and retained. The major importance of genetic potentialities is implied by our recognition of the need for children to have reached a certain stage of maturation or reading readiness before it is profitable to embark on formal reading. Out of school experience, encouragement or discouragement at home, physical health and temperamental adjustment are also very influential, as Burt's and Schonell's investigations of backwardness have demonstrated. Thus, it is clearly false to think of school attainment as consisting simply in the absorption of information and skills that the school provides.

A further reason for regarding intelligence as a kind of attainment, and attainment as a kind of intelligence, is that there seems to be little difference in their stability or variability over periods of time. Children do, of course, vary from class to class in, say, their progress in arithmetic, but their all-round level appears to remain fairly steady in a majority, just as does the I.Q., except perhaps when they change school or take up very different school subjects.

Nevertheless, a partial distinction still seems useful. Intelligence B refers to the more general thinking capacities—comprehending, problem-solving, rational thinking, level of concept development, etc., which have very wide applicability or transfer-value to daily-life situations as well as to educational materials or to intelligence-test questions. Whereas attainments refer to more specific capacities which are more directly moulded by stimuli received in the classroom, and which probably depend to a somewhat greater extent on interest and on personality factors such as perseverance and motivation. But the distinction is a matter of degree rather than of kind, and there are many intermediate capacities which are difficult to classify. For example, vocabulary, general knowledge, and mathematical reasoning are sometimes included in intelligence, sometimes in attainments, tests.

The main corollary of this viewpoint is the abandonment of the conventional conception that intelligence is a cause of attainment, and that backwardness can be explained away if it is found that a child has a low I.Q. It is true that both Intelligence B and any kind of attainment are largely dependent on, or caused by, Intelligence A, but as we have no means of assessing Intelligence A accurately, that does not take us any further. In the majority of children, I.Q. and average educational quotient go pretty closely together; the correlation is about 0.85. But quite a number can be found in whom one is appreciably higher or lower than the other. Surely, this should not surprise us any more than does the observation that some are better at verbal abilities, others at mathematical, or at practical, abilities. There may be certain differences in inherited talents; but apart from these we would expect different interests and experiences to produce stronger or weaker development in some abilities than in others.

The Achievement Quotient

This brings us to the topic of the Achievement Quotient, which has aroused so much controversy recently—that is the ratio of Educational to Mental Age, or Educational Quotient to I.Q. As is well-known, it is supposed to show whether children are "working up to their native capacity." If not, they are called " retarded," as distinct from merely backward, and educational psychologists or remedial teachers tend to explain the discrepancy in terms of environmental, temperamental or other factors, and to regard such pupils as the most fruitful cases for psychological or remedial educational treatment. The discussion centred around the explanation of A.Q.s above 100, i.e. children who were working " above capacity."

But there was little consideration of the points dealt with in this article—points which imply that the whole notion of working at the level expected from the I.Q. is unsound. In addition, as I have shown elsewhere, the A.Q. should be abandoned since it is a highly unsatisfactory type of measurement. It is apt to be very unreliable or untrustworthy; and inevitably it is unfair to the brighter child. Far more above than below average I.Q. children are likely to come out as retarded, although, on psychological grounds, it is the below average ones who are more likely to be failing to realise their potentialities.

How, then, should we interpret the more extreme discrepancies? When E.Q. greatly exceeds I.Q. it may be that the child is unusually docile, persevering, interested in school work, and well coached at school or at home, but is not advancing as rapidly in general understanding, nor applying his attainments to non-scholastic problems. Alternatively it may be simply that the particular kind of intelligence test does not suit him; it may be highly speeded and he prefers to work slowly, or if it is non-verbal he may lack the spatial or other ability that it involves over and above *g*.

The opposite type of discrepancy also frequently arises because the intelligence test embodies speed, spatial or other group factors in which the child is relatively strong, but which have scarcely any bearing on his educability. Nevertheless, as we know from follow-up studies of the 11 +, a reliable verbal intelligence test does provide somewhat better evidence of future educability than does present attainment, when the circumstances of schooling alter. Thus, a proportion of children with E.Q.s much below I.Q.s probably does arise from environmental handicaps, emotional maladjustment, lack of interest, irregular schooling, and the like. Though these same factors are liable to depress the I.Q. also, they may have less effect on thinking abilities which are built up to a greater extent from home and leisure experiences.

A general conclusion that follows is that intelligence tests are somewhat less valuable in educational guidance and in the diagnosis of backwardness than earlier psychologists supposed. On the other hand, the views expressed here do nothing to lessen their value in educational selection and prediction. General thinking skills as well as attainment in junior school subjects are highly relevant to grammar school success. Similarly, in streaming, or in allocating the very dull to special classes, a combination of intelligence and attainments tests is likely to be more prognostic than either separately.

Which children, then, may we regard as failing to ' work to capacity '? According to the view of this article, all children are underfunctioning to a greater or lesser degree, in the sense that neither their attainments nor their general intellectual development are as advanced as they might be, given ideal home upbringing and schooling. Social and educational policy have done much to improve this,* but the problem is still so vast that we are forced to concentrate our remedial efforts on those most likely to benefit. Now, in general, the most serious cases are those whose attainments lag furthest behind their chronological ages.

But, in addition, the skilled teacher or the educational psychologist can often discover more hope of improvement in some very backward children than in others. Unfortunately, we have to admit nowadays that an I.Q. on a verbal test, while of some relevance in picking out these cases, is less helpful than was previously supposed; and a non-verbal test is seldom of any use at all. Research evidence gives very little support for the claim that higher I.Q. children overcome their ' retardation ' more rapidly than low I.Q. children; and it has not yet revealed any more scientific means of diagnosis than the teacher's or the clinic's subjective judgment. Of course, the child with I.Q. 120, E.Q. 100 often appears to respond more readily to treatment than another of the same age with I.Q. 70, E.Q. 80, because he is much less seriously backward and easier to make contact with. It might be argued, too, that he is more worth treating as being potentially a more valuable member of the community. Yet, so far as our present knowledge goes, the latter might gain as much as, or more than, the former from careful diagnosis of his difficulties and individual attention.

One other practical point to emphasise is that non-verbal or pictorial group tests, and individual performance tests, are much inferior to verbal ones in predicting educability (except, possibly, in predicting ability for science and technical courses among older pupils). Thus, their use for scholastic purposes should be avoided. They may appear to give a fairer chance to very young children, or to older pupils who are poor readers. But with few exceptions (such as testing the deaf), they do not yield a better index of ' innate potentiality '. Group oral tests should, therefore, be used when individual Binet testing is impracticable.

---

*Thus, in the United States the average army recruit in World War II scored the equivalent of some 10 I.Q. points higher than in World War I. While this may partly reflect the greater familiarity of intelligence tests in the 1940's, it is probably mainly due to nation-wide improvements in the general educational level.

References for Further Reading

ANASTASI, A., *Differential Psychology* (rev. edit.). New York: Macmillan, 1958.

CLARKE, A. D. B., *et al.*, " Cognitive and Social Changes in the Feeble-minded "—Three Further Studies. *Brit. J. Psych.*, 1958, XLIX, 144-157.

HEBB, D. O., *The Organization of Behaviour*. New York: Wiley, 1949.

O'CONNOR, N. and TIZARD, J., *The Social Problem of Mental Deficiency*. London: Pergamon Press, 1956.

PIAGET, J., *The Psychology of Intelligence*. London: Routledge & Kegan Paul, 1950.

PIDGEON, D. A., YATES, A., BURT, C., *et. al.* The Relationship between Ability and Attainment. *Bull. Nat. Found. Educ. Res.*, 1956-7. Nos. 8-10.

VERNON, P. E., " The Relation of Intelligence to Educational Backwardness." *Educ. Rev.*, 1958, XI.

VERNON, P. E. (edit.), *Secondary School Selection*. London: Methuen, 1957.

YATES, A. and PIDGEON, D., *Admission to Grammar Schools*. London: Newnes, 1957.

TRANSFER AT ELEVEN PLUS—

A Summary of Evidence provided by Research

by ALFRED YATES and D. A. PIDGEON

Introduction

THE arrangements made by local education authorities for the transfer of children from primary to secondary schools are carefully and sometimes anxiously scrutinised by teachers, parents, and members of the general public. The decision to transfer a child to a particular type of secondary school not only determines the kind of curriculum that he will pursue but may also affect his chances of proceeding to some forms of higher education, and help to define his social status and the range of his vocational opportunities. For these reasons the eleven plus is generally regarded as a process of selection and popular appraisal of the methods used by local education authorities tends to be based on other than purely educational grounds. In spite of the efforts made by authorities to represent and indeed to treat transfer at eleven plus as one phase in a continuous process of educational guidance, it has become a question in which social and political considerations are involved, with the result that public discussions of the problem are often emotionally charged and there is a tendency to confuse a number of issues that in the interests of clarity should be kept distinct.

It is legitimate, for example, to ask each of the following questions: to what extent are current methods successful in attempting to discriminate between children suitable for grammar, technical and modern schools or courses as these are organised at present?; how far does the provision of schools or courses of these kinds satisfy the educational needs of the children concerned?; is it in the best interests of society as a whole to organise secondary education in this way? The first of these questions poses an essentially technical problem. The second question concerns the relationship between, on the one hand, the abilities, attainments and interests of primary school leavers and on the other, the opportunities afforded for their development by existing secondary school courses. The third embraces the more comprehensive problem of determining the needs of modern society and of examining the adequacy of our present educational arrangements to satisfy them.

In current discussions of the problems of transfer from primary to secondary schools, arguments relevant to the consideration of one of these questions are often used to answer another. The logical error that occurs most frequently consists in levelling criticisms against procedures of selection that are in fact based on objections to the ends that these procedures are designed to serve. Some parents view with disquiet the fact that what appears to them to be a largely irreversible decision affecting their children's educational and vocational prospects is made when the children are ten years of age. The target for their criticism, however, is often the examination—or some component of it—on which the decision is based.

We have provided this introductory comment on the circumstances in which judgment tends to be formed of the efficiency of the arrangements for transfer at eleven plus in order to define the scope of this article. The research evidence available is almost wholly concerned with the first of the three questions listed above. In the nature of things it is existing arrangements rather than proposed reforms that can be subjected to rigorously controlled inquiry. The evidence that we propose to summarise, therefore, is derived from researches that have examined the operation of arrangements for allocating children to secondary school courses within what has been termed the tripartite system. Since there have been very few researches of a large-scale or intensive kind into selection for technical courses the bulk of the evidence is, therefore, related to the problem of how most effectively to discriminate between those children who appear to be suitable for a grammar school course and those who do not. It is only to this narrow and specific question that an answer based on a wealth of research evidence can be provided. The broader questions concerning the kinds of secondary school organisation that could best serve present-day needs have not yet been systematically examined.

In the summary that follows we have not, except in a few instances, referred to the specific researches that have contributed to the evidence, nor entered into any discussion of the design and conduct of the inquiries or the methods of analysis used. For those who wish to consult the evidence at first hand we have provided a bibliography containing a list of the sources consulted.

The contents of the eleven plus examination

We propose to use the term ' eleven plus examination ', in accordance with popular usage, as an umbrella term to cover all the procedures—whether these consist in fact of written examinations and tests or of interviews, teachers' assessments or other devices—which education authorities employ as a basis for their allocation of pupils to different types of secondary school course.

Since each authority is free to devise its own form of examination and there is, consequently, a large variety of procedures in use, it is imprecise to speak of ' the ' eleven plus examination. This variety, however, consists for the most part in different combinations of a relatively small number of possible components and in different ways of treating the results that they yield, and so by considering the evidence relating to each of the parts separately and to the more commonly employed combinations, a reasonable estimate can be formed of the efficiency of the procedures as a whole.

1. *Intelligence tests*

Pride of place must go to the intelligence test both because this is the feature of the eleven plus examination which has occasioned a good deal of controversy and because its use, in spite of the criticisms levelled against it, is widespread. (At the time of the last survey carried out by the Foundation in 1956 it was found that just over 88 per cent. of local education authorities were using intelligence tests as part of their eleven plus procedures.) Numerous researches have shown that intelligence tests predict success in grammar schools more closely than any other type of written test or examination that has as yet been tried.

In spite of the fact that these results have been widely publicised there is still a considerable amount of opposition to the use of such tests. Part of this undoubtedly stems from the mistaken notion that they yield, in some mystical fashion, a measure of innate ability. It is far less wounding to his pride for a parent to be informed that his child's attainments in arithmetic are below a required standard—this can be attributed to misfortune or bad teaching—than to be told that his child has failed to satisfy the examiners in a test which is popularly supposed to reflect his degree of inherited ability. Neither of the two organisations—the National Foundation and Moray House—that supply the bulk of the intelligence tests set by local education authorities uses the word intelligence in the title of its tests. Moray House has recently changed the description to ' Verbal Reasoning '. The Foundation, from the outset, avoided any trace of the tendentious by styling them simply as ' Verbal ' tests. It is unlikely, however, that the popular use of the term ' intelligence tests ' as a description of these instruments will disappear for some time and the associated idea that they involve the measurement of some basic and largely inherited quality will also probably linger.

A somewhat more serious objection to the use of intelligence tests in the eleven plus examination is that research has shown that they are susceptible to the effects of coaching and practice. As a consequence many children are involved, in or out of school, in preparatory activities which take up time that could perhaps be more profitably spent in other pursuits. This objection, probably as a result of the prejudice against the supposed nature of intelligence tests has been considerably exaggerated. Indeed, this charge is sometimes brought against intelligence tests without acknowledgement of the proven fact that all the other kinds of test or examination that are or can be used in the eleven plus examination are equally affected by special preparation.

As far as the predictive value of intelligence tests is concerned, coaching is only likely seriously to distort the outcome if some candidates have been expertly prepared for the tests whilst others are totally unfamiliar with them. This situation is hardly likely to arise nowadays. The fact that the effects of coaching are minimal unless the material used for the purpose is similar in content and format to that included in the test proper has been amply demonstrated. Moreover, it has also been shown that the advantages accruing to those who have been expertly coached can be offset by arranging for all children who are about to enter for the examination to be allowed opportunities to familiarise themselves with the type of test to be encountered.

The only valid objection remaining against the use of intelligence tests, therefore, is the fact that if they are used in the authority's examination they are likely to be introduced into the curricula of some primary schools. This is tantamount to saying that they have what has come to be called a ' backwash ' effect on the work of the primary schools. Again this is a criticism that can be levelled against other tests or examinations. Whatever form of examination is used at the end of an educational course, it will tend to influence the content of the curriculum and the methods of teaching employed.

In deciding whether or not intelligence tests should be used in an eleven plus examination, therefore, the major consideration should be whether or not the particular kind of backwash effect associated with these tests is sufficiently serious to warrant the exclusion from the examination of one of its most efficient components. We believe that some, at least, of the activities in which primary school children are involved when they are being coached for an intelligence test—discovering synonyms, opposites and analogies, doing simple exercises in verbal reasoning, breaking codes, analysing number series, etc.—would find a place in the curriculum, even if intelligence tests were no longer included in the eleven plus examination.

2. *Tests of attainment*

The volume of popular criticism levelled against the use of intelligence tests has tended to obscure the fact that tests of attainment in English and arithmetic are used in most examinations, and since equal weight is usually given to each of the three parts of the examination, a child's level of attainment is a more decisive factor than his performance in the intelligence test. The majority of authorities employ standardised objective tests of attainment, but there has been a tendency in recent years for the reintroduction of the older forms of examination, in English particularly, sometimes in place of standardised tests but more often in addition to them. The reasons for this apparent ambivalence—the wish to restore the older form of examination and yet to retain the newer type of objective test—are first, that considerable dissatisfaction has been voiced, particularly by those who teach in grammar schools, about the effects on the attainments of primary school children of the exclusive use of objective tests; and, secondly, that objective tests have been demonstrated to be both more reliable[1] and more valid[2] than their predecessors.

The arguments against the use of objective tests are familiar enough. They refer to the 'backwash' effect of tests which rely in the main on 'multiple choice' items—questions for which a series of possible answers are supplied, the examinee being required to indicate the one acceptably correct response. It is often alleged that if such tests are used to assess attainment in English, children in primary schools tend to be deprived of adequate opportunities to practise writing in consecutive prose. Although evidence to support this charge is somewhat meagre it has, nevertheless, achieved a widespread currency.

There is considerable evidence relating to the relative efficiency of objective tests and the older forms of examination. A follow-up study* carried out in the area of an authority which employs both types of examination showed that the prediction afforded by the objective tests alone was as high as that for the total examination. This result can be justifiably interpreted in two ways. Since objective tests alone achieved as good a result as a procedure involving the addition of older forms of examination, the latter could be dispensed with, thus reducing the amount of time and money spent on the process. Alternatively, since the addition of the older forms of examination did not *reduce* the efficiency of the examination, they may safely be included in order to offset the supposedly adverse 'backwash' effects of the objective tests.

An attempt has been made to devise a new kind of English test that makes the best of both worlds, so to speak. Such tests, although constructed on the pattern of modern objective tests in that they contain considerably more items than the traditional type of examination, nevertheless follow the example of the latter in requiring creative responses, some including the construction of complete sentences. It may be supposed, therefore, that the 'backwash' effect of these new type tests is not dissimilar to that of the older examinations, yet they have been shown to be as reliable as the purely objective tests and to enjoy even a slightly higher level of predictive efficiency.*

---

[1] The reliability of a test is the degree to which it is likely to yield consistent results on successive occasions.

[2] The validity of a test is the extent to which it achieves its purpose—i.e. in this instance, affords a prediction of success in a grammar school.

*See Yates and Pidgeon (1957).

English essays have also been reintroduced into the procedures used by many authorities in recent years. The reason for doing so is, no doubt, to encourage those primary school teachers who, in preparing their pupils to face objective tests, are alleged to follow an unduly restrictive curriculum. The researches of Wiseman (1949) and others have encouraged authorities to take this step, by demonstrating that it is possible to employ methods of marking essays which yield satisfactorily reliable results.

3. *Teachers' assessments*

It can scarcely ever have been doubted that the teachers who have had charge of a child's primary education are able to make a sound appraisal of his level of ability and attainment and can, therefore, predict with some accuracy his chances of success in a grammar school course. The major problem has been that of discovering a satisfactory means of collating the assessments provided by teachers from different schools. It has been demonstrated that teachers can satisfactorily rank their own pupils with respect to their suitability for a grammar school course.

It is clearly important, however, in awarding grammar school places to children from different primary schools, to make due allowance for the fact that the proportion of suitable candidates in one school is not necessarily the same as that in another. Various devices, designed to make this necessary adjustment, have been tried. The quota scheme, pioneered by the Walsall authority, the methods of scaling and quantifying teachers' assessments introduced by Sutcliffe and Canham (1945) and by the National Foundation (Yates and Pidgeon, 1957), and the procedure devised by G. F. Peaker and tried out in part of the West Riding, represent attempts to place on a comparable basis assessments made by teachers in different schools. All but the last of these makes an adjustment based on test scores or examination marks. The West Riding procedure relies on the observed stability of the proportion of successful candidates in each of the primary schools concerned.

There is a considerable body of evidence to show that if teachers' assessments are suitably adjusted in one of the ways described above they yield a predictive measure equal to that provided by an examination and, indeed, superior to that provided by any single test. This fact has raised the important question as to whether it might be possible to dispense with an eleven plus examination and to rely wholly on the assessments of their pupils provided by primary school teachers. The available evidence suggests that the combination of teachers' assessments and a soundly constructed examination is more efficient than either used alone. It may well be possible, however, by improving the methods whereby teachers' assessments are made and utilised to reach a level of predictive efficiency that could only be improved marginally, if at all, by the addition of an examination.

---

*Op. cit.

Some take the view that the use of teachers' assessments, without an examination, would be desirable even if there were some consequent loss of efficiency (the evidence suggests that this would be slight) on the grounds that primary schools would be liberated from the restrictive influences of the examination. Against this it is necessary to set two considerations: in some areas, particularly those in which competition for grammar school places is severe, teachers may not be willing to shoulder the whole burden of allocating pupils to different types of school; the public, generally, has a higher regard for procedures of selection that are based on apparently impartial examinations than for those which are more directly dependent on human judgment.

The problem of the borderzone

No known form of examination or assessment has achieved perfect reliability. The practical consequence of this is that if any eleven plus examination were to be repeated—the tests given on a different day, for example, or the teachers' assessments requested in a different form—there would be some inevitable variation in the results. In other words, drawing a line across a list, on whatever basis the list has been compiled, in order to distinguish between those to be admitted to a grammar school and those to be excluded cannot be justified. For this reason it is generally recognised that the dividing line should be replaced by two lines, dividing the candidates into three groups: those about whose suitability for a grammar school course there can be no reasonable doubt; those who are clearly unsuitable; and that group of candidates, usually described as belonging to the borderzone, about whom no firm prediction can be made on the basis of the evidence available.

The problem of delimiting the border-zone is one that can be satisfactorily solved by discovering empirically either the degree to which the tests and assessments used fall short of perfect reliability, or the accuracy with which they predict future performance. The problem of what further assessments to make of the potentialities of the candidates within it in order to arrive at a more satisfactory prediction is far less tractable. The law of diminishing returns stubbornly frustrates attempts to achieve the fine discriminations that are required at this stage in the selection procedure. One self-evident principle, which is nevertheless frequently ignored, is that in dealing with border-zone candidates techniques of assessment should be sought which are more comprehensive and effective than those which have achieved the initial relatively coarse groupings. It is scarcely reasonable to expect, for example, that a brief interview can enable discriminations to be made which a combination of examination scores and assessments provided by teachers and based on continuous and close observation over a period of years have failed to achieve.

The evidence suggests that to improve on the prediction afforded by a sound initial procedure—based, for example, on suitably scaled teachers' assessments and a battery of standardised tests—requires the collection of a considerable amount of additional evidence concerning the child's abilities and attainments and a procedure whereby this evidence can be reviewed globally by an experienced panel of judges. Even an elaborate and time-consuming procedure of this kind has been found to yield only a small improvement.

The effects of differences in age

There is evidence to show that differences in age, even those amounting to no more than a few months, affect the performance of children in examinations and tests of all kinds. Thus, if no allowance were made for these differences, there would be a tendency for the older children within the age group considered for admission to grammar schools, to gain a disproportionate share of the available places. The use of standardised, objective tests, counteracts this tendency in that appropriate allowances for differences in age are incorporated in the system of transforming the ' raw ' or actual marks gained by a child into the standardised scores which are used to express his relative level of performance.

Other forms of assessment, however, are not necessarily free from these effects. Many of the authorities that use unstandardised examinations make some adjustments to the scores to offset the injustice that would otherwise be done to the younger candidates. These adjustments, however, are not adequate in every instance and, moreover, it is rare to find that age allowances are considered in the case of English essays, primary heads' assessments, or interviews. Since primary heads' assessments have been shown to be of considerable value as predictors of academic success and are for this reason becoming more widely employed it is important that their main defect—the demonstrable tendency for these assessments to be biassed in favour of older children—should be remedied. A number of methods have been devised for this purpose and are described in detail in Yates and Pidgeon (1957) and Vernon (1957).

Differences in test scores between boys and girls

The task of authorities in attempting to administer an eleven plus examination that is manifestly fair to all candidates is further complicated by the fact, established by evidence from a number of sources, that the performance of girls is superior to that of boys in most if not all of the tests and examinations commonly employed. This superiority of girls appears to be maintained for some years after the children have been transferred to secondary schools. It is reasonable to suppose that it is associated with the fact that girls mature at an earlier age than boys and that therefore the differences manifest at eleven may well disappear subsequently, although there is, as yet, insufficient evidence to allow this conclusion to be drawn with certainty. In the circumstances the most satisfactory course for an authority to adopt would seem to be that of treating boys and girls separately for the purposes of allocation to secondary school courses.

The validity of the examination

The most crucial question to ask about the eleven plus examination is to what extent it serves its major purpose of distinguishing between those who are and those who are not suitable for a grammar school course. It is impossible, in the nature of things, to answer this question directly and empirically. To do so it would be necessary not only to discover a generally accepted criterion of success in a grammar school course but also to arrange for all children to follow such a course. Only by this means would it be possible to obtain direct evidence concerning the suitability or otherwise of those candidates excluded by the selection procedure. Another complication, although one that can be to some extent resolved, arises from the fact that there are marked differences between schools of the same type, so that it may well be that a child who does not satisfy the requirements of one grammar school, for example, may well have prospered had he been allocated to another.

Three main criteria have been employed by those who have attempted to evaluate the effectiveness of the eleven plus examination: G.C.E. or School Certificate results; internal school examination marks; and the assessments provided by the secondary school teachers concerned. It may well be argued that any one of these criteria, or any combination of them constitutes an inadequate measure of the extent to which children have been appropriately allocated. On the other hand, an eleven plus examination which failed to discriminate between those children on the one hand who are able to put up a satisfactory performance in school examinations and to earn the commendation of those who teach and those, on the other, who are adjudged incapable of reaching the standards required of them could scarcely be defended. Criteria depending on examination marks and teachers' assessments of their pupils' suitability for a grammar school course may not be wholly adequate for the purpose but there do not appear to be any practicable alternatives.

The researches carried out by the Foundation used the two latter criteria, and by comparing them, for each secondary school separately, with the results of the eleven plus examination, it was possible to assess the degree to which the latter had afforded an adequate prediction. The analyses were carried out for the schools separately to allow for the fact that there are marked differences between schools even of the same type.

It is clear that if all the comparisons showed perfect agreement—that is, if the eleven plus examination placed children in exactly the same order of merit as that which was subsequently found to obtain in the secondary school to which they were transferred, it would be reasonable to assume that the children ' above the line ' on the original examination list were, in fact, more worthy of admission to grammar schools than those who fell beneath it. On the other hand, it also follows that if a lack of perfect correspondence is discovered then there must have been some error in the initial allocation, and the extent of this error can be estimated by observing the degree of correlation between the eleven plus results and the subsequent performance of the children concerned.

The Foundation's estimate of the size of this error—an estimate which is in agreement with that made by other workers [see Vernon (1957)]—was that, with the best available procedure, some ten per cent. of the candidates are likely to be ' wrongly ' allocated. ' Wrongly ' is placed in inverted commas because this statement does not necessarily imply that, as a result of this apparent error, the children concerned are likely to receive educational treatment inappropriate to their needs. The differences between children who form the border-zone are often very small. The child who is allocated to a secondary modern school, but who is nevertheless slightly superior academically to one who just manages to secure admission to a grammar school, is not necessarily deprived of opportunities to develop his potentialities. Nor is it by any means certain that his slightly weaker colleague will derive no benefit from his experience in a grammar school.

On the other hand, since errors of this order are likely to be made by even the most efficient procedures that have as yet been devised, the policy, now adopted by more than half the education authorities, of so modifying the organisation of secondary education as to blur the hitherto sharp distinctions between grammar schools and other types of secondary school would seem to have much to commend it.

There is little doubt that the eleven plus examination, as developed by many authorities, is the most efficient examination to be found in our educational system. It is almost certainly more reliable than any of the examinations that children encounter at later stages in their careers and, as a process of selection, it is probably more valid than those employed for the selection of entrants to Universities, the Civil Service, or the various professions. The main criticism that can justifiably be brought against it is that it has at times been required to accomplish the virtually impossible task of segregating into courses which are sharply differentiated, children between whose abilities, attainments and ascertainable educational needs there is a scarcely perceptible difference.

Select bibliography

BOSOMWORTH, G., "The Use of Teachers' Assessments in Allocating Children to Secondary Schools." *Durham Res. Rev.*, Vol. IV, 1953.

BURT, C., "Validating Tests for Personnel Selection." *Brit. J. Psych.*, Vol. XXXIV, 1943.

CLARK, H., "The Effect of a Candidate's Age upon Teachers' Estimates and Upon His Chances of Gaining a Grammar School Place." *Brit. J. Ed. Psych.*, Vol. XXVI, 1956.

DEMPSTER, J. J. B., *Selection for Secondary Education.* London: Methuen, 1955.

EMMETT, W. G., *An Inquiry into the Prediction of Secondary School Success.* London: University of London Press, 1942.

EMMETT, W. G., and WILMUT, F. S., "The Prediction of School Certificate Performance in Specific Subjects." *Brit. J. Ed. Psych.*, Vol. XXII, 1952.

EMMETT, W. G., "Secondary Modern and Grammar School Performance Predicted by Tests Given in Primary School." *Brit. J. Ed. Psych.*, Vol. XXIV, 1954.

FLOUD, J. E., *Social Class and Educational Opportunity.* London: Heinemann, 1957.

FREEMAN, F. S., and MILES, C. C., "Sex Differences." *Encyclopedia of Educational Research*, Rev. Ed. London: Macmillan, 1950.

HARTOG, P. J., and RHODES, E. C., *An Examination of Examinations.* London: Macmillan, 1935.

HARTOG, P., RHODES, E. C., and BURT, C., *The Marks of Examiners.* London: Macmillan, 1936.

JAMES, W. S., *et al.*, *Studies in Selection Techniques for Admission to Grammar Schools.* Bristol Univ. Inst. Educ. Publ. No. 3. London: Univ. of London Press, 1952.

McCLELLAND, W., *Selection for Secondary Education*, 1st Imp. London: Univ. of London Press, 1942.

McINTOSH, D. M., *Promotion from Primary to Secondary Education.* London: Univ. of London Press, 1948.

McINTOSH, D. W., WALKER, D. A., and MACKAY, D., *The Scaling of Teachers' Marks and Estimates.* Edinburgh: Oliver & Boyd, 1949.

MINISTRY OF EDUCATION, *Early Leaving: A Report of the Advisory Council for Education (England).* London: H.M.S.O., 1954.

NATIONAL UNION OF TEACHERS, *Transfer from Primary to Secondary School.* London: Evans Bros., 1949.

PAULEY, F. R., "Sex Differences and Legal School Entrance Age." *J. Ed. Res.*, Vol. XLV, No. 1, 1951.

PEEL, E. A., and RUTTER, D., "The Predictive Value of the Entrance Examination as Judged by the School Certificate Examination." *Brit. J. Ed. Psych.*, Vol. XXI, 1951.

PILLINER, A. E. G., "The Position and Size of the Border-Line Group in an Examination." *Brit. J. Ed. Psych.*, Vol. XX, 1950.

RICHARDSON, S. C., "Some Evidence Relating to the Validity of Selection for Grammar Schools." *Brit. J. Ed. Psych.*, Vol. XXVI, 1956.

RUTTER, D., "An Inquiry into the Predictive Value of the Grammar School Entrance Examinations." *Durham Res. Rev.*, Vol. I, 1950.

SUTCLIFFE, A., and CANHAM, J. W., *Selection for Secondary Education.* London: John Murray, 1945.

Times Educational Supplement, "Selection at Eleven—III. Tyranny of Tests. Yorkshire System of Exemption." 23rd September, 1955.

VALENTINE, C. W., and EMMETT, W. G., *The Reliability of Examinations.* London: Univ. of London Press, 1932.

BUREAU OF EDNL. & PSYL. RESEARCH LIBRARY 1-9-70 8257

VERNON, P. E. (Ed.), *Secondary School Selection*. London: Methuen, 1957.
WALKER, A. S., *Pupils' School Records*. London: Newnes, 1955.
WATTS, A. F., and SLATER, P., *The Allocation of Primary School Leavers to Courses of Secondary Education*. First Interim Report. London: Newnes, 1950.
WATTS, A. F., PIDGEON, D. A., and YATES, A., *Secondary School Entrance Examinations*. Second Interim Report. London: Newnes, 1952.
WIGFORSS, FRITS, *The Entrance Examination in View of Later School Performances*. Stockholm: P. A. Norstedt & Soners Förlag, 1937.
WISEMAN, S., " The Marking of English Compositions in Grammar School Selection." *Brit. J. Ed. Psych.*, Vol. XIX, Part III, 1949.
WISEMAN, S., *Problems Connected with the Selection of Children for Secondary Schools*. Ph.D. Thesis, University of Manchester, 1952.
WRIGLEY, J., " The Relative Efficiency of Intelligence and Attainment Tests as Predictors of Success in Grammar Schools." *Brit. J. Ed. Psych.*, Vol. XXV, 1955.
YATES, A., and PIDGEON, D. A., *Admission to Grammar Schools*. Third Interim Report. London: Newnes, 1957.

THE WISH TO LEARN

Research into Motivation

BY W. D. WALL

I. Introductory

The Will to Learn

THE teacher is directly and continuously concerned with learning and with the will to learn. He frequently observes that a child who finds it genuinely difficult to remember a minimum of facts necessary, say, to the comprehension of geography will commit to memory without effort equally difficult facts concerning aircraft or ' pop ' singers. He probably notices, too, that inside and outside the classroom, his pupils acquire a good deal of knowledge, a great many techniques, and attitudes which are only incidentally related to what they have been specifically taught at home and at school or even to what they may themselves have deliberately set out to acquire.

Educationally, we are faced with a major problem. What is this will to learn and how do we direct it to those bodies of knowledge, those techniques and attitudes which it is the business of the teacher to impart? How can we see that it leads to incidental learning of the right kind? What is it in the original learning situation, or subsequently, that leads to retention of what has been learned? What is it in one piece of learning that facilitates or hinders the pupil in his subsequent learning?

Primary Motivations.

All this directly concerns what the psychologist calls ' motivation '—perhaps the most obscure and difficult of all the theoretical and practical issues in general and educational psychology. It does, however, seem to be very clear that the concept of motivation covers a variety of aspects of human (and animal) behaviour and that, for clarity of thinking and for practical use, it is as well to make a few distinctions at the outset.

Some motivations are evidently primary—that is to say, that they spring from more or less crude biological drives like hunger, thirst, sex and so on. In their primitive forms, they certainly directly and powerfully motivate human behaviour and the more they are denied, the greater becomes the tension in the organism until they are satisfied. It is certain, too, that a drive like hunger will lead an animal or a human being to learn very rapidly the ways by which it can be appeased and that such learning is likely to be durable.

The primary drives of hunger and thirst seem to be the main springs of learning in the first few days and weeks of life; but even they rapidly become, under normal circumstances, somewhat sophisticated and overlaid with other learnings and developments of ' secondary' or ' derived ' motivations. More and more, as the child grows older, these secondary motivations become important and it is difficult to see a primary drive as the mainspring of any learning which he may do in school. Indeed, we are often tempted to see such primary motivations as hindering factors. Hunger, fear, aggressiveness, the search for security and love, if they reach any level of intensity in a child of school age may well lead to behaviour and learnings which run counter to what the teacher is trying to teach or what the parents wish from their child. So far as, for example, learning to read or to calculate are concerned, the appeal is to some such motives as interest, self-assertion through success, affection for the teacher and the like which are derived or secondary, that is, they are themselves learned and elaborated out of previous learning experience.

Unconscious Motives

We have so far spoken as if the will to learn were always conscious. This, however, is far from being the case, even with the intelligent and sophisticated adult. It becomes painfully apparent when we try to teach a child who has, for example, developed an emotional difficulty and fear of arithmetic. He may consciously wish to succeed; and he may try and try hard: but he evidently has an unconscious counter motivation which defeats his efforts and ours. Similarly, there is little doubt that children and adolescents learn rapidly and apparently without conscious effort many difficult things which satisfy needs which they could not formulate consciously. In this form of learning we may surmise, from work on child development and from the studies of maladjusted children, that such basic drives as the search for affection and acceptance, the need for security, the desire to assert oneself and the sex-drive, particularly in adolescence, are of primary importance.

Incentives

Clearly, there are some circumstances in which these inner motivations, whether they are conscious or not, are sufficient in themselves to lead to action, to seeking behaviour the end of which is to satisfy the need; and clearly, if the environment prohibits certain easy ways of satisfying the need, the child (or the animal) will learn the relatively more difficult techniques fairly rapidly.

It is on this basis that much of the early learning and training of infants takes place. The infant seeks to satisfy its hunger; the mother provides the satisfaction but at the same time deflects the child from one form of behaviour to another. Similarly, the infant seeks the warmth and physical security which nursing provides and mother tends to use this as a reward for steps forward in behaviour and self-control which the child learns. Thus, from a very early age, parents, and particularly mothers, tend to use *incentives:* that is, the infant is stimulated and motivation is aroused by something extrinsic to himself. This, however, as every mother quickly finds out, will only be successful if the incentive used really does stimulate some need in the child and if there is no competing need of greater strength.

In the class-room, this interaction between more or less primary motivations and the incentives which the teacher offers, is even more complex than in the early years of life and the interaction is less evidently direct. We may illustrate this by the frequently recurring situation in which the teacher offers as an incentive praise which is to be earned for work well and successfully achieved. Such praise means, to the five-year-old child, say, affection and the security of being loved. Thus, the anticipation of it is a stimulus to his search for acceptance and security. Thus far the situation is likely to arouse effort to do what the teacher wishes him to do.

But there may be a competing need. The child may be afraid of the teacher; or he may have sat still already for so long and the need for physical movement is greater than anything else. The incentive of praise will not work if other motives are operating more powerfully. Nor will it work if the task which has to be undertaken is beyond the child's capacity. As we shall see later, if strong motivation is aroused in the face of an impossible task, behaviour may become more disorganised in proportion to the strength of the individual drive. Indeed, we may say that conflicting drives and a disproportion between the motivation aroused and the possibility of achievement are the two circumstances which do most to disorganise learning in children and which are most likely to lead, if they are repeated, to serious difficulties in growth.

Group Influences

In very early life, it seems likely that all motives are individual and personal and probably in the early school years this may continue to be largely true. However, children very rapidly come to form a part of the class group and of sub-groups within the class or within the school. All teachers sooner or later become aware of the fact that there is such a thing as group motivation. The stimulus which the teacher provides to learn may well operate differently according as to whether it is given in a direct individual context or whether it is given to the class as a whole. Not infrequently, to our regret, we find that the class or a section of it produces a powerful counter-motivation to what we are trying to do; and there is a kind of solidarity which no pupil seems willing to break.

Conversely, if once a group accepts and makes its own a motive put forward by the teacher, the learning which is undertaken will be strikingly effective even for those members of the group who, individually, would find the work beyond their powers. It seems probable that, in the junior school and onward until well into adolescence, this social or group motivation, as we may call it, is more important educationally with normal children than is anything more evidently individual and direct.

Clearly, motivation in human learning is a highly complex thing, difficult to study objectively, and instinct theory, in its earlier form, even when elaborated by secondary or derived motivations, does not suffice to account for everything, even in the learning of children. There is a further important notion, first formulated by Allport, which we must add and it may be that it is, educationally, the most significant of all. While it is true that much human behaviour can be accounted for as the result of biological drives of an innate, more or less instinctive kind, and of their elaborations and derivatives acquired through the process of living and learning, it is also true, as Allport points out, that some motives may be directly acquired and achieve an autonomy of their own. An obvious example is smoking. Once the habit is acquired, it will lead to all kinds of behaviour and may have much the same motive force behind it as hunger. It is a part at least of the task of education at home and at school to create desirable motives and habits which become 'functionally autonomous'.

II. Research Work

The purpose of this somewhat lengthy introduction is to make clear some at least of the complexities of the subject under study and to suggest a part at least of the reason why much of the research evidence is difficult to interpret and to apply in the day-to-day teaching situation. In spite of the apparent paucity of direct classroom studies, however, there is much to be gleaned from the literature of research both in general and in educational psychology.

Animal Learning

Since the thirties, and particularly since the early work of Thorndike, there has been a great deal of work on the learning of animals and a consequent exploration of the effectiveness of what we have called earlier, primary motivations, like hunger, thirst, sex and the like. The general upshot of this work is, as we might expect, a confirmation that such primary motivations are extremely powerful.

It has also become clear that they are by no means the sole motivations seen in animal learning nor are they necessarily effective in the same way at different levels of intensity. It is apparent that too great an intensity is likely to lead to disorganisation of behaviour which in some cases looks like panic.

It has, too, become clear that motivation is a blend of two things: a state of inner tension which may be high or low; and the external stimulus which may be strong or weak. Motivation depends upon the relationship between the two at any given moment. It is held, too, that, in general, animal learning is the result of conditioning: that is to say, that a successful response is learned because it reduces the need or motivation; and an unsuccessful response is eliminated because it does not lead to the satisfaction of a need. However, while this is generally true, certain experiments have shown that ‘ punishment ’ (i.e. usually electric shock in these experiments) can and does lead to learning.

Curiosity and Manipulation

Crudely, these findings concerning primary motivations apply to human learning, especially in the learning of babies; but they do not suffice to account for learning in the school situation or indeed for most of what occurs after the first few months of life, particularly in the intellectual field. There are two other findings from the work with animals which are more directly relevant. Some at least of the studies of ape learning have demonstrated that there are motivations other than the primary drives or needs. When these primary drives are at a low level, i.e. when the animal is satisfied and not fatigued, something which might be described as ‘ curiosity ’ occurs and the animal will explore more or less at random. Similarly, too, apes will show ‘ pleasure in manipulation ’—they will repeat successful activity with puzzle apparatus not now for a direct reward but apparently for the sheer pleasure of doing it.

Observations of pre-school children tend to suggest that ‘ curiosity ’ and ‘ pleasure in manipulation ’ are two of the most important factors in learning and in the mastery of skills or knowledge, which may have been acquired under other stimuli—and this even from the earliest months of life. Babies tend to play with a skill once acquired. Finally, it has been shown in a number of experiments with animals that there is such a thing as incidental learning—that is, learning which is apparently unmotivated—though it is difficult to demonstrate absence of motivation in animals and *a fortiori* in humans.

Insight and Feed Back

In most of the conditioning or learning experiments undertaken with animals the process exhibited by the animal is the laborious one of 'trial and error'. The rat running in the maze, the monkey in a puzzle cage, continue more or less at random to try to get at the bait, slowly eliminating errors, but often without that leap forward which comes when insight occurs and enables it to short cut the trial process. Experiments with human maze learning and some experiments with animals have shown, however, that insight does occur and that some learning is not simply conditioning in the old trial and error sense. Evidently in the teaching situation 'insight'—which may be the pupil's own more or less unaided discovery of a short cut to the solution of a problem or the result of skilful teaching—is of primary importance. In the context of the study of motivation it is equally so since the taste of sudden success and the insight itself tend not only to reinforce the original drive but may provide a notice on their own account and stimulate further effort.

Animal psychology can throw no light on certain further aspects of our topic which are of fundamental importance. Unconscious motivations, the relationships which may well exist between motivation and incentive on the one hand and the subject's level of aspiration—that is to say, the goal which he sets himself—on the other, and the generally greater complexity and indirectness of human functioning, make direct inferences from animal studies to human psychology hazardous.

Basic Needs of the Personality

Certain things are, however, clear. The concept of basic drives and of 'need reduction' is of the greatest utility when we are considering infants and children of pre-school age; and certainly much if not all early learning is motivated in this way. However, from the outset of life sheer manipulation and 'curiosity' are manifested by babies whose other needs are satisfied and the early stages of play as a means of learning and expression and of the construction of symbolic thought are assigned by Piaget to this.

From the point of view of the child of school age, however, the principal fact of importance is that curiosity, interest and the desire to manipulate and explore can hardly manifest themselves unless the basic needs of children are more or less continuously satisfied. It goes without saying that such things as hunger, fatigue, and marked insecurity are liable to inhibit forms of activity which are not directed to their removal or satisfaction; and we know from several generations of work with maladjusted children that a denial of the need for affection, of acceptance, of the need for an outlet for aggression will sooner or later lead to a failure in school learning. It is frequently not so clearly seen that such basic needs of the personality are for normally developing children an important part of the whole context of school learning, often in ways which are difficult to analyse and which differ somewhat from age to age.

When the child begins school it seems likely that he views his teacher in much the same way and with something of the same attitudes as he does his mother or father. The praise and blame of the infant mistress, her acceptance or rejection, are important motives which will in large part determine his behaviour and his eagerness to learn. It seems probable, from the numerous studies which have been made of teacher characteristics and of the relation between liking for a teacher and liking for the subject which he teaches that such direct personal interaction of pupil and teacher remains an important source of motivation at least until well on into adolescence.

Nevertheless, the normal progression of growth should be that, because the basic needs of the personality are satisfied in the child's home life and generally in school, his energy is steadily freed to become occupied with the objective work. This 'liberation of the self from the self', to quote Hamley's phrase, can only proceed steadily provided that outlets, of a legitimate and constructive kind, are offered for aggression and self-assertion and provided, too, that home and school environment generally favour increasingly social and heterocentric attitudes. In other words, school learning depends mainly upon secondary and derived motivations which can only be powerful stimuli of use to the teacher if the pupils are satisfied in other more basic directions.

In this, certain broad rhythms of development must be taken into account. In early school years primary motivations—love, aggression, fear—may tend to predominate in many children, though side by side with them we can see evidence of much curiosity and activity for its own sake. Normally, in the junior school period, the emphasis seems to shift to secondary motivations; children enjoy acquiring skills and knowledge, particularly if their desire for mastery and achievement is fostered by praise and success. Later, the onset of puberty and adolescence seems, for many, to sharpen the demands of primary needs, particularly self-assertion, sex and the desire for security, this time of an emotional—social kind.

It is often difficult for the teacher of young adolescents to stimulate them to learning of the conventional type; their interests and their needs lie elsewhere. It is questionable whether it is wise to persist in somewhat arid reliance in abstract curiosity, disinterested desire for achievement and the like in such cases. Probably a recognition of the pre-occupations of the growing boy and girl, of their personal and social problems, of their fears and anxieties (many of them homely enough and clumsily formulated if at all) and the adaptation of curricula and method to them would produce more effective and rapid learning.

In the second half of the 'teens' usually motivation appears to change. By this time many boys and girls have left school and the realities of life appear to provide acceptable stimuli to effort. Not infrequently, the 'lazy' and 'apathetic' youth of 15 turns into the young man or woman who is working hard at evening school or technical college or catching up his apparently lost years.

It is against this broad background that we should consider and interpret the evidence which we have as to the effectiveness of the various 'secondary' or 'acquired' motivations.

Curiosity and Interest

It will be recalled that the work with animals has shown that purely manipulative and exploratory activity and apparent curiosity occur even in rats and more so in monkeys. It seems, too, that the amount of this as well as its apparent force as a motive, increases as one goes up the scale in species and within species according to intelligence. Experienced teachers are well aware that the more intelligent children are likely to display more interest and more curiosity than their less able contemporaries. Equally, such intelligent children are likely to have a wider and more mature range of interests. The problem is to know the precise psychological function of curiosity and interest and the means of rousing and sustaining them as motives for educational purposes.

There are a great many studies of the interests of children and adolescents at various ages which tend to show the existence of general tendencies in the development of interest, some sex differences in interests especially in the ' teens ', and a wide range of individual variation. When the declared interests of a given child or group of children are closely examined one tends to find that the overt content corresponds to some deeper and more or less permanent personal need which may find expression in many different ways at different times. Similarly, the same expressed interest may mean very different things in different children of the same sex and age. Hence, the advice often given to teachers to make use of the interests of pupils as a motivation to school learning needs some qualification and interpretation.

The work of Gardner, Daniels and others in England, and of the Ecole Nouvelle (Freinet, Cousinet and others) abroad, though little of it is in the nature of rigorously controlled experiment, nevertheless shows how effective the expressed interests of children may be to secure initial motivation for a task. Such work also, however, shows that the child's interest and curiosity alone cannot be relied on to carry him through the learning process without other stimulation and encouragement. It would seem that the teacher's task is to accept the declared interests and curiosity of his pupils as a starting point and indeed to use them as far as possible all along the way. At the same time, however, she must so arrange the environment that one interest provokes another and that the feeding of interest or the satisfaction of curiosity are made sufficiently easy so that the child can with reasonable effort experience success, fairly continuously.

That this is so is abundantly clear from the detailed information given by Susan Isaacs in her *Intellectual Development of Young Children*. The environment of the Malting House School was so arranged as to contain a great deal of material and the adults helped the children consistently to find solutions to the problems which they set for themselves. The children asked questions and initiated activities; but so also did the adults. The art consisted in initiating an interest or an activity which seemed to be already latent in the group. The role of the teacher, therefore, is certainly to guide, through the provision of materials, through directing attention to problems and processes likely to interest children, through asking appropriate questions and through tactful help at critical points in finding solutions or in carrying out tasks which may be within the conception of the pupils but beyond their immediate ability to execute.

Praise and Blame

The maintenance of interest and curiosity, however, at least in part, depends upon success and acceptance and praise that effort brings. Some light is thrown on this by an early experiment by Hurlock. She took four groups of children, all of whom were set the task of addition. Initially, all groups achieved approximately twelve points of score on the test used. One group was then praised for its performance; a second was reproved for the mistakes which it had made; a third was ignored, although the pupils heard the praise and reproof meted out to the others; finally, there was a control group. After a period of this treatment, the four groups were retested with the results that: the praised group had a mean of twenty points; the reproved group fourteen points; the ignored group twelve points; and the control group eleven points.

This experiment suggests that praise is a powerful stimulus to effort at least with primary school children. It also seems to be clear that blame or punishment of whatever kind has the effect of inhibiting a particular form of action without necessarily putting anything positive in its place. The effects of blame seem to be primarily emotional, affecting performance rather than learning and in situations where what is wanted is unclear or beyond the capacity of the pupil it may lead to the repetition of undesirable behaviour or to confusion. The relationship of this last observation to some of the more stubborn cases of arithmetic or reading failure is evident.

Lewin also points out that a reward or punishment situation, particularly where it is directed to individuals in the class group, must imply some form of policing; that is to say, that barriers are erected against the attainment of the reward by a short cut such as copying the answers. Such externally imposed frustrations are bound, he suggests, to have a profound effect upon the social group concerned.

The effectiveness of praise or blame seems to be conditioned by other factors. A research by Thompson and Hannicutt with pupils in the Vth grade (i.e. about 11-12 years old) suggested that repeated praise tended to increase the work of those children who were more introverted while the more extroverted children were more stimulated by blame.

Still more is it likely to be conditioned by the prior experience of pupils, especially in so far as they have been successful or unsuccessful in similar activities before. Three researches are of interest here, carried out by Sears, by Gebhard and by Cartwright. Sears found that pupils with a prior experience of success tend to be realistic in their goal-setting in arithmetic and reading, that is to say that their level of aspiration more or less coincides with their possibilities. On the other hand those who have experienced failure in similar tasks, either set unrealistically high goals or they tend to state a much lower level of aspiration as a kind of self-protection against further failure. Cartwright and Gebhard, working independently, discovered that success and the anticipation of success increased considerably the attractiveness of otherwise neutral tasks.

The effects on children of praise and blame, success and failure are thus difficult to predict in any given case. It is evident that failure in a general context of success has a different meaning from a failure which is one more of a long chain. It is evident, too, that the degree to which the child feels himself to be generally acceptable to his teacher will condition markedly the effect of blame or failure on his attitudes to further effort.

Anxiety

There is some evidence that a certain amount of anxiety in otherwise normal children provides a motivation for learning. Robertson, in an unpublished research on children whose performances were better or worse in the secondary grammar school than was predicted by the eleven plus allocation tests, finds some support for the idea that children who improve are those who are mildly anxious about their work. Too great an anxiety on the other hand (whether about school matters or about matters in the out-of-school life), may lead to inhibition of effort or passivity. Equally, it seems that children who are without anxiety do not do as well as might be expected. This last finding coincides with other work that suggests that the ablest children in a group tend to conform to the norm rather than too markedly to excel. Lynn, studying the reading and arithmetic attainments of primary and secondary school children, found that anxiety tended to be associated with a higher level in reading than in arithmetic in otherwise normal groups.

A level of anxiety which is noticeable to the teacher may, however, tend to inhibit and disorganise effort—particularly effort in such subjects as arithmetic which depend upon undisturbed concentration. Many normal educational circumstances, notably examinations, provoke evident anxiety in children and adolescents and this may in fact adversely affect performance even of normal pupils. Calvin, McGuignan and Sullivan, working with undergraduate women students, found that examination performance was markedly improved if an outlet for such anxiety (in the form of free comment) was permitted.

Work of this kind, and general psychological considerations, suggest that the teacher should stimulate children to effort and achievement—perhaps rather more than is customary in some progressive schools and activity classes—but that such stimulation should be adjusted by an understanding of the individual pupil's general proneness to anxiety and insecurity. Furthermore, it seems likely that freedom on the part of the pupils to express their anxieties coupled with skilful reassurance on the part of the teacher is likely to ensure that such anxiety is in fact creative rather than inhibiting.

Group Influences

So far, we have spoken of the child as an individual or as one side of the teacher-pupil relationship; he is, however, much more than this. At least when he comes to school and often before, he becomes aware of the group of his contemporaries and equals and slowly begins to develop a group loyalty. Indeed, the feeling of belonging either to a small group of friends or to a larger class-group is an important element in his feeling of security; and remaining acceptable to the group becomes itself an important motivation for him. Within the group, too, he has a role to sustain which is partly a matter of his own choice and partly what the others reflect upon him. Thus, by middle childhood at least, what might be called socially determined motivations are of great importance inside and outside the classroom and the simple, direct relationship of teacher and taught no longer quite holds. The aims of the teacher and those of the taught may be in conflict without either side being fully aware of why, or even that this is so.

Some of the work which has been done on the relationships between the personality and methods of the teacher and the reactions within the class is highly relevant here. Anderson and his collaborators, working with nursery school classes, found that the teacher who tended to dominate provoked resistance, fidgeting and over-conformity in her pupils. In contrast the teacher who attempted integrative responses to children and who sought to draw them into activities with their own free consent was rewarded by more spontaneity and more initiative in learning from her class. Lewin, Lippit and White, working with somewhat older children and with adolescents in youth clubs, contrasted three broad kinds of approach: dictatorial; laissez-faire; and democratic leadership. They found that hostile and aggressive reactions were as much as thirty times more frequent in autocratic groups. In the groups which were democratically led there was higher morale, more communication, closer friendships, less discontent and aggressive behaviour, and more work-minded conversation. Similar conclusions are reached by Richardson in her studies of teacher behaviour and adolescent reactions.

Co-operation and Competition

Closely related to this are experiments which have been conducted to determine whether co-operative working or interpersonal competition gives the better results. Deutsch, in a study of two matched groups of school children, found that competition resulted in much more destructive and aggressive criticism while the co-operative group showed more co-ordination of effort, greater pressure to achieve, more productivity per unit of time, and a higher quality of the end-product. In addition there were advantages of a social kind shown in friendliness and attentiveness to others.

Thelen, summarising the upshot of a long series of researches carried out in the classroom laboratory of the University of Chicago, puts forward ' some significant propositions ' as follows: (*a*) the teacher's behaviour in a large measure determines the quality of the emotional conditions in the classroom; (*b*) the learning of certain social attitudes and human relations principles is affected by teacher-pupil interaction; (*c*) the teacher-pupil interaction patterns may affect the pupil at deep (i.e. subconscious) levels; (*d*) pupil-pupil interaction (for example, status-roles) can be affected by the nature of the teacher-pupil interaction. He further points out that much of this can be generalised from school to adult groups; conversely, of course, much of the work on decision-making in industry is applicable to school situations and it has been over and over again found that a genuinely democratic participating group tends to produce in each member more satisfaction in the task undertaken and a tendency to retain the group decision as one's own.

A word of caution, however, is necessary in attempting to apply these findings. It does seem clear that in general the best motivation is obtained when the teacher involves the class in the making of the decisions as to what shall be done and how. Perhaps the best and almost the sole large-scale demonstration of the effectiveness of this is the well-known Eight Year Study in which the best results were obtained from those schools which approached their work in such a way as to obtain the maximum participation of the students.

However, although accepting and democratic, the teacher should not abandon his necessary role of leadership. He has a contribution to make to the decision-making and to the choice and setting of aims which springs from his superior knowledge and experience. He must, too, allow for the fact that not all pupils will have the same willingness or ability to participate and that some at least may well be isolates who do not respond at all or at least directly or desirably to group influences. He has to guard against the tendency of the abler members to adapt their effort and pace to that of the group rather than to exert themselves to the full—thus, he may have in fact to combat an attitude, itself the product of group loyalty, that the individual should not differ too much from his companions. Equally, he has to be aware of the tendency of some children to become passengers.

He should, too, know that he cannot change the climate of a group abruptly from the more usual one of more or less benevolent autocracy to a truly democratic one; the change has to be mediated and the necessary attitudes have in fact to be learned by all who participate, including, very often, the teacher himself. Finally, the nature of the leadership which can be given by the teacher will depend considerably upon the degree of maturity of the pupils and upon the realism with which they can set goals for themselves. In general, the older and more able the group of pupils, the more fully democratic the leadership can be.

Socio-economic Status

We have so far spoken as though the motivations for learning were either individual or a function of the nature and climate of the classroom group. This, of course, is only partly true. Outside these and beyond the control of the teacher lie certain social motivations, difficult to analyse, but which have to be taken into account. Our knowledge of these is sketchy and indirect, and the interpretation of the data which we have is somewhat speculative. There is some evidence, both from English and from American sources, to suggest that general interest and success in school work are positively related to socio-economic status, even when general ability is held constant. The interest of the parents in their children—an influence beginning well before school years—their own experience with schooling and their degree of education seem also to be directly connected with levels of attainment. On the other hand, there may be, in our community, a rural group and a small urban one also which is indifferent if not hostile to education.

Attitudes and Opinions

The influence of the social group operates in other ways. It is well known that the attitudes and opinions of children and adolescents tend to be closer to those of their parents than to those of their teachers and that the vocational choices which are made by boys and by girls tend in many cases to be related to the occupation of the parents. An enquiry by Jahoda carries this into more detail. He studied school-leavers from secondary modern schools and found that whatever their actual social class membership, boys tended to identify themselves with the skilled manual and routine grades of non-manual work. They intended to ‘ stay with the lads ’ and declared that they wished to become fully skilled craftsmen, which was what a ‘ job with prospects ’ implied.

The girls tended to be less realistic in their assessment of the present status of their parents and in their ambitions; many of them, too, seemed to express a conflict between their desire to move up in social grade and loyalty to the group to which they are attached. In the ‘ teens ’, and especially in the last year or two of compulsory schooling, such vocational identifications may be important sources of motivation; but they do not necessarily operate in the direction proposed by the teacher.

C

Conclusion

It is apparent from the analysis we have made of some of the work on children's motivations that no simple rules and precepts can be given to teachers. It is, too, very clear that psychologists are far from having unravelled the whole delicate system which determines whether or not a class will put forth a maximum effort in school. However, such knowledge as we do have can lead us to a number of practical inferences which, if they guide classroom practice, are likely to be helpful.

The first of these is embodied in the proposition that if the fundamental needs of the growing personality are not fully met, there will be little energy left for intellectual things. Some, even most, of the dissatisfaction and maladjustments of children and adolescents arise outside the school. This, however, does not mean that the school is powerless to do anything about them. Severely maladjusted children should, of course, be referred for examination and treatment to the school psychological service; but all children at some stage show signs of strains in their growth which may affect their effort in class. It is then that they specially demand perceptive handling and support from their teachers and parents.

With advancing maturity and emotional balance, children normally learn to differentiate the demand for love which they make on their parents from that they make of their teachers and friends. But in spite of his deceptively self-contained air the primary school child still needs affection and acceptance in school; and so less obviously perhaps but none the less imperiously does the adolescent. An impersonal or worse still, a hostile attitude on the part of the teacher, or the school staff generally, is not conducive to effort at any stage. Warmth and understanding from the teacher are essential to good pupil-teacher relationships and it is in the context of these that praise and blame, punishment, reward and criticism must be viewed as stimuli to effort.

Children of all ages have other needs which should not be denied. Some outlet for aggression, some chance of physical movement, the security which comes from an orderly and comprehensible framework of do's and don'ts and from a reasonable but not obsessive rhythm to the school day, are likely to allow the secondary motivations of curiosity and the desire to explore and manipulate to emerge. So, too, self-assertion, both as an individual and as member of a group or team, calls for the acceptable and constructive outlet of success and a place in the sun.

Play and fun, too, are fundamental to growth and not merely to the young child. Through play and fantasy children and adolescents can come to grips with and solve some of their problems. Many children in our society do not have sufficient opportunity to play in their out-of-school lives and suffer in consequence. A major psychological argument for carefully planned active methods is that among other things, they allow for those who are deprived of play.

Some work by Ojemann and by Selley has drawn attention to the influence on children's school progress of greater knowledge by their teachers of childhood and its problems and of the value to normal school work of the chance given to primary school pupils of discussing their personal problems freely. This does not imply an airy-fairy 'psychiatrisation' of the school. It suggests that when outlets are given for pressing needs and when the teachers thoroughly understand the emotional growth of children, then their disciplinary systems, and the incentives which they use to increase motivation are more likely to be effective and less likely to run counter to the deep and powerful drives and preoccupations of their pupils.

Fear and anxiety have been shown to be, on the whole, inhibitive and disorganizing factors in learning. Some children it is true do seem to thrive under punishment, but in the long run, it seems more likely to lead to undesirable forms of social learning—cheating, deception, and so on—than to any good effects. This, however, does not mean that punishment, reproof and blame should never be used. In the context of healthy pupil-teacher relationships and when they are clearly deserved, they are likely to be effective as deterrents. However, this effect is largely lost if at the same time pupils are not constructively shown the way to win praise and approval.

It is evident, too, that children's levels of aspiration, the goals they set themselves, are related to two things which are within the control of the school. Prior knowledge and learning and especially the experience of successful effort, affect incidental learning and present motivation. Thus, the adjustment of the gradient of difficulty in all tasks which pupils undertake is vital not merely for the matter in hand but for the growth of attitudes towards work and learning in the future. This does not mean that work should be so arranged that success comes easily. It means that the demands made are such as to evoke the maximum effort and that the teacher should be watchful to see that failure is only a rare occurrence; he should intervene to help when the child is clearly struggling with something beyond his powers or will to succeed. Here, too, we should point out the value of positive habits of getting down to a task in the expectation that an honest effort will bring success or at least understanding from the teacher.

Finally, the importance of the class as a group and of the sub-groups within the class must not be overlooked and it is in drawing attention to this that perhaps the greatest contribution has been made by recent research. The system of discipline used by the teacher, his personality and the relationships he builds with individual pupils all affect the climate and consequently the motivation of the class. He is most likely to succeed if he becomes aware of the feelings of group solidarity which exist, if he is perceptive of those children who tend to be isolated and if he attempts democratically to ensure that the pupils participate in the determination of the immediate, and with older children and adolescents, the more remote aims of their own education. This does not mean that he abrogates his duty to give direction, to set standards, to maintain discipline; it means that he exercises it in a way which is intelligible and acceptable to his pupils, and in a framework in which each child is sure of being cared for. Even more important perhaps is that in such a social climate he is more likely to be able realistically to assess what his pupils can achieve and help them, through the experience of success, to develop the habit and the will to learn.

Bibliography

This is a list of the material consulted and summarized in the preceding paper. Items providing a select initial bibliography for further reading are marked with one asterisk (*); more difficult and specialised items with two asterisks (**).

ANDERSON, H. H., *et al.*, " Studies of Teachers' Classroom Personalities." *Appl. Psychol. Monogr.*, 1945-6, Nos. 6, 8, 11.

AUBLE, D., and MECH, E., "Response Strength in a Classroom Task Related to a 'Forward' Delay in Reinforcement." *J. Ed. Psych.*, Vol. XLV, No. 3, March, 1954.

BROWN, W. F., ABELES, N., and ISCOE, I., "Motivational Differences between High and Low Scholarship College Students." *J. Ed. Psych.*, Vol. XLV, No. 4, April, 1954.

*BUGELSKI, B. R., *The Psychology of Learning*, Chap. 9, "Motivation and Learning." London: Methuen, 1956.

CALVIN, A. D., McGUIGAN, F. J., and SULLIVAN, M. W., "A Further Investigation of the Relationship between Anxiety and Classroom Examination Performance." *J. Ed. Psych.*, Vol. XLVIII, No. 4, April, 1957.

CAMPBELL, W. J., "The Influence of Home Environment on the Educational Progress of Selective Secondary School Children." *Brit. J. Ed. Psych.*, Vol. XXII, Part II, June, 1952.

*CARROLL, HERBERT A., "Motivation and Learning: Their Significance in a Mental-Health Program for Education." *Mental Health in Education*, 54th Yearbook of the U.S.A. National Society for the Study of Education, Part II. Univ. of Chicago Press, 1955.

CARTWRIGHT, D., "The Effect of Interruption, Completion and Failure upon the Attractiveness of Activities." *J. Exp. Psych.*, Vol. XXXII, 1942.

CASSELL, RUSSELL N., "Motivation as a Synthesis of Contemporary Psychology." *J. Ed. Psych.*, Vol. XLIII, No. 3, March, 1952.

COLLINS, M., "The Causes of Premature Leaving from Grammar Schools, Part II." *Brit. J. Ed. Psych.*, Vol. XXV, Part I, February, 1955.

CRAIG, R. C., "Directed versus Independent Discovery of Established Relations." *J. Ed. Psych.*, Vol. XLVII, No. 4, April, 1956.

*DANIEL, M. V., *Activity in the Primary School*, Oxford, Blackwell, 1947.

DEUTSCH, M., "An Experimental Study of the Effects of Co-operation and Competition upon Group Process." *Human Relations*, Vol. II, No. 2, 1949.

DREYER, A. S., "Aspiration Behaviour as Influenced by Expectation and Group Comparison." *Human Relations*, Vol. VII, No. 2, 1954.

**EYSENCK, H. J., *The Structure of Human Personality*, Chap. VI, "The Analysis of Physiological Measures" London: Methuen, 1953.

*FLEMING, C. M., "The Child Within the Group—The Bearing of Field Theory and Sociometry on Children's Classroom Behaviour." *Studies in Education, No. 7: The Bearings of Recent Advances in Psychology on Educational Problems*. London: Evans Bros., 1955.

*GARDNER, D. E. M., *Long Term Results of Infant School Methods*. London: Methuen, 1950.

GARDNER, D. E. M., *Testing Results in the Infant School*. London: Methuen, 1942.

GEBHARD, M. E., "The Effect of Success and Failure upon the Attractiveness of Activities as a Function of Experience." *J. Exp. Psych.*, Vol. XXXVIII, 1948.

HARLOW, HARRY F., "Motivational Forces Underlying Learning." In *Learning Theory, Personality Theory and Clinical Research*. The Kentucky Symposium. Lonson: Chapman & Hall, 1954.

**HEBB, D. O., *The Organization of Behavior*, Chap. 8, "Problems of Motivation—Pain and Hunger." London: Chapman & Hall, 1949.

HIERONYMUS, A. N., "A Study of Social Class Motivation: Relationships between Anxiety for Education and Certain Socio-Economic and Intellectual Variables." *J. Ed. Psych.*, Vol. XLII, No. 4, April, 1951.

*HIGHFIELD, M. E., and PINSENT, A., *A Survey of Rewards and Punishments in Schools*. Publication No. 3 of the National Foundation for Educational Research. London: Newnes, 1952.

HILGARD, ERNEST R., "Human Motives and the Concept of the Self." *The Study of Personality*. London: Chapman & Hall, 1954.

*HILGARD, E. R., and RUSSELL, D. H., "Motivation in School Learning"; SWENSON, E. J., "Applications of Learning Principles to the Improvement of Teaching in the Early Elementary Grades"; THELEN, H. A., and TYLER, R. W., "Implications for Improving Instruction in the High School." *Learning and Instruction*, 49th Yearbook of the U.S.A. National Society for the Study of Education, Part I. Univ. of Chicago Press, 1950.

HUGHES, E. W., "Children's Choices in Individual Activities in the Junior School." *Brit. J. Ed. Psych.*, Vol. XXV, Part I, February, 1955.

HURLOCK, E. B., "An Evaluation of Certain Incentives Used in School Work." *J. Ed. Psych.*, Vol. XVI, 1925.

**ISAACS, S., *Intellectual Growth in Young Children*. London: George Routledge & Sons, 1930.

JAHODA, G., "Social Class Attitudes and Levels of Occupational Aspiration in Secondary Modern School Leavers." *Brit. J. Psych.*, Vol. XLIV, Part 2, May, 1953.

KEMP, LESLIE C. D., "Environmental and Other Characteristics Determining Attainment in Primary Schools." *Brit. J. Ed. Psych.*, Vol. XXV, Part II, June, 1955.

KITTELL, JACK E., "An Experimental Study of the Effect of External Direction during Learning on Transfer and Retention of Principles." *J. Ed. Psych.*, Vol. XLVIII, No. 7, November, 1957.

*KNIGHT, REX, "Children's Needs and Interests—Contemporary Theories of Motivation." *Studies in Education, No. 7: The Bearings of Recent Advances in Psychology on Educational Problems*. London: Evans Bros., 1955.

*LAIRD, A. J., and KNIGHT, A. R., "Contemporary Studies of Motivation." In *Current Trends in British Psychology*. London: Methuen, 1953.

LEWIN, K., "Field Theory and Learning." In *The Psychology of Learning*, 41st Yearbook of the U.S.A. National Society for the Study of Education, Part II. Univ. of Chicago Press, 1942.

LIPPIT, R., and WHITE, R. K., "The Social Climate of Children's Groups." Chapter XVIII in Barker, Komim and Wright (Ed.), *Child Behaviour and Development*. New York: McGraw-Hill, 1943.

Lynn, R., " Temperamental Characteristics Related to Disparity of Attainment in Reading and Arithmetic.' *Brit. J. Ed. Psych.*, Vol. XXVII, Part I, February, 1957.

McGeogh, J. A., and Irion, A. L., " Learning as a Function of Motive-Incentive Conditions." In *The Psychology of Human Learning*. New York: Longmans, Green & Co., 1952.

Morgan, M. I., and Ojemann, R. H., " The Effect of a Learning Programme Designed to Assist Youth in an Understanding of Behaviour and its Development." *Child Development*, Vol. 13, No. 3, 1942.

*Morris, Ben, " Mental Health in the Classroom." *Studies in Education, No. 7: The Bearings of Recent Advances in Psychology on Educational Problems*. London: Evans Bros., 1955.

*Oeser, O. A. (Ed.), *Teacher, Pupil and Task—Elements of Social Psychology applied to Education*. London: Tavistock Publications, 1955.

Ojemann, R. H., and Wilkinson, F. R., " The Effect on Pupil Growth of an Increase in Teachers' Understanding of Pupil Behaviour." *J. Exp. Educ.*, Vol. 8, December, 1939.

Radke-Yarrow, M., and Yarrow, L. J., " Child Psychology." *Annual Review of Psychology*. Stanford, California-Annual Reviews Inc., Vol. 6, 1955.

**Richardson, J. E., *et al.*, *Studies in the Social Psychology of Adolescence*. London: Routledge and Kegan Paul, 1951.

Sawin, E. I., " Application of Learning Theory to Appraisal of Motivation: A Pilot Study." *J. Educ. Res.*, Vol. XLIX, No. 8, April, 1956.

Sears, P. S., " Levels of Aspiration in Academically Successful and Unsuccessful School Children. *J. Abn. and Soc. Psych.*, Vol. XXXV, 1940.

Seeley, J. R., *Crestwood Heights*. Toronto: 1956.

**Sherif, M., and Cantril, H., *The Psychology of Ego-Involvement*. London: Chapman & Hall, 1947.

Silberman, H. F., " Effects of Praise and Reproof on Reading Growth in a Non-Laboratory Classroom Setting." *J. Educ. Psych.*, Vol. XLVIII, No. 4, April, 1957.

Simpson, R. H., " An Evaluation of Motivation, Assignment Making, Problem Procedures, and ' Service Learnings ' in a School System." *J. Educ. Res.*, Vol. XLIV, No. 1, September, 1950.

Stiles, F. S., " Developing an Understanding of Behaviour at the Elementary School Level." *J. Educ. Res.*, Vol. XLIII, 1950.

Thelen, Withall, Flanders, Rehage, Perkins, Glidewell, Symposium on " Experimental Research Towards a Theory of Instruction." *J. Educ. Res.*, Vol. XLV, No. 2, October, 1951.

Thompson, C. G., and Hunnicutt, C. W., " Effect of Repeated Praise or Blame on the work achievements of Introverts and Extroverts." *J. Educ. Psych.*, Vol. XXXV, May, 1944.

Teaching Children to Read—I

THE RELATIVE EFFECTIVENESS OF DIFFERENT METHODS OF TEACHING

A. *The Place and Value of Phonics*

by JOYCE M. MORRIS

THE effectiveness of any teaching method is determined by school conditions as well as by the attributes of teacher and pupil. It is therefore unlikely that research will ever reveal one best method of teaching reading but, when sufficiently developed, it may help the inexperienced teacher to choose suitable reading methods in a given set of circumstances, and thereby avoid much wasteful experimentation in the classroom.

Basic Methods of Teaching Reading

Today selection may be made from a great number of reading systems of which almost all are modifications and combinations of a few basic methods. Two of these—the alphabet-spelling method which originated in ancient times, and the look-and-say or word method devised by Comenius in the 17th century, were almost universally employed in our early infant schools. By these methods children learning to read the word CAT, for example, either named the separate letters as SEE AY TEE and followed this by saying CAT, or were taught to recognise the word correctly at sight without more ado.

The alphabetic method is undoubtedly of value in learning to spell but of limited usefulness for the reading of words and sentences; the look-and-say method, although based on the psychologically correct idea that a word is not merely a collection of letters but has a recognisable character or pattern of its own, does not by itself give pupils a means of tackling new words independently. Hence, it is not surprising that these early methods lost ground rapidly when phonic methods were widely introduced and developed. With the substitution of sounds for letter-names and their subsequent blending to form a whole word, the word CAT, for instance, could now be attacked as follows: KUH-A-TT, CAT. This type of blending method is still used, especially by the older generation of teachers, but in modern phonic teaching it is more usual to find children being taught to divide simple words into initial consonants and phonograms. Thus, the word LAND would be divided into L-AND, and then by combining the phonogram with different preceding and following letters a word family may be built up, e.g., band, stand, sandy, etc. Training in the recognition of common syllables and in analysing long words by syllables is now a recognised part of phonic training.

Phonic methods help children to learn to read in a systematic left-to-right direction and provide them with a technique by which unfamiliar words may be deciphered. However, apart from the fact that a considerable proportion of English words are non-phonetic, other disadvantages of these methods are frequently quoted. It is claimed that the mechanical nature of any phonic method lays too little stress on the importance of comprehension in reading, and that the practice of splitting up words into their common sounds or combinations of sound hinders whole word recognition. The phonic method imposes restrictions on the vocabularies of reading primers so that they are often poor in content and artificial in style. Habits of slow, laboured reading are also said to be developed by this method.

Different phonic systems flourished at the beginning of this century, but, when the limitations of a purely phonic approach were realised, teachers turned again to the older methods. The highly mechanical, difficult and uninteresting alphabetic method was never reinstated to its former position in the reading programme, but the value of a sight vocabulary was again recognised with the result that a 'mixed' method was recommended whereby regular words were taught phonically and irregular words by look-and-say.

Fresh support was also given to advocates of 'analytic' as opposed to 'synthetic' methods when the sentence method of teaching reading was popularised in the thirties. In this method (which is claimed to be a good one because it follows natural lines), the sentence, rather than the word, is regarded as the unit of meaning and emphasis is laid on reading comprehension. Its great value lies in the help offered by the context which, although limited to a certain extent by the need for frequent vocabulary repetition, does not have to be of the cat-sat-on-a-mat variety. Children taught exclusively in this way, however, may be satisfied with guesses instead of looking carefully at individual words, and they will probably be bewildered when they meet a word which is new to them. Difficulty in distinguishing between words of similar size and shape may also be experienced.

Since satisfactory progress is not guaranteed by the use of these visual and auditory techniques alone, many teachers include kinaesthetic methods in their reading programmes. These methods, which stress the importance of tracing and writing in learning to read, were first described over two thousand years ago. More recently, Montessori employed them, but they were not recognised widely as basic procedures until 1921 when Fernald and Keller reported their favourable effects on the development of word recognition in non-readers.

Kinaesthetic methods reinforce other sensory impressions of words, develop a consistent left-to-right direction in reading, and provide adequate drill and repetition. They also encourage careful and systematic observation and enable both teacher and pupil to measure progress easily. Three major criticisms of the exclusive use of these methods may be offered. First, they do not give pupils a means of deciphering new words without assistance. Secondly, they develop certain movements of arm, hand and lips which may interfere with later progress in speed and fluency. Finally, the necessarily individualised instruction which these methods entail make them generally more suitable for small classes.

There are thus advantages and disadvantages in all the main methods of teaching reading used in the past. Consequently, most teachers today do not confine themselves to one particular method, but, recognising the importance of individual differences, provide sufficient variety in their approach to reading to enable their pupils to read by the look of a word, by its sound, by guessing from the context, and by utilising the sense of movement. They differ, however, in the proportions of each method which make up their basic reading programme, and in the extent to which they adjust this proportion to the needs of individual children. There are also differences in the emphasis which teachers place on oral and silent reading as well as in the amount of systematic as opposed to informal instruction given. It is at this point that controversies arise, and where scientific investigation may be of practical use in deciding what is empirically sound educational practice and what is merely a matter of ' fashion '.

Reading Methods and Research

Since the second decade of this century experiments to determine the relative effectiveness of different teaching methods have been carried out. Unfortunately, few investigators have exercised sufficient care in ensuring that differences in results have been due to teaching method and not to the skill of the teachers, the attributes of the children or other school characteristics. Methods, too, have not been clearly defined, and uniformity of name does not necessarily guarantee uniformity of method. Thus, seemingly conflicting evidence has been produced on many controversial questions.

Several attempts to draw together what is certainly known about reading in general have been made, among the most notable being those of Gray, Vernon, and Traxler and Townsend. Recent events, however, have focussed attention specifically on reading methods, and on the value of phonic as opposed to whole-word methods in particular. In 1954, public concern about school methods was aroused in America by the publication of a book by Flesch addressed to parents and entitled *Why Johnny Can't Read*. The author claimed that there had been a deterioration in children's reading standards which was largely due to the neglect of phonic teaching in the modern school. Flesch also maintained that current educational practice is not in line with scientific evidence, since he can find none in favour of the whole-word approach although he is able to quote a number of studies which prove the superiority of the phonic method. A similar conclusion was reached by Daniels and Diack in this country, for they stated that: "In all the experiments reported in the literature, we know of only three which have not come to the definite conclusion that some variety of the phonic method gives better results than a whole-word method as far as word recognition is concerned." According to these researchers even this small amount of evidence is suspect owing to the experimental design of the three investigations cited.

These statements and the concern of parents and teachers naturally call for a re-examination of modern methods of teaching reading and the scientific bases on which they rest. Since, however, the literature on reading methods is extensive, the present discussion will be confined to one main topic only, namely, the *place and value of phonics in the reading programme*.

Consideration of what research tells us about phonics necessarily includes references to other methods, but the major purpose of this first article is to review studies which help to answer such questions as: Should phonics be taught? When and how should phonic training be given? Is a particular kind of reading skill produced by systematic phonic instruction? Are phonic methods more effective with certain types of children, e.g., the dull and backward? First it is necessary to meet the challenge that phonics teaching is neglected in the schools of today.

The Place of Phonic Methods in Current Educational Practice

Before the book by Flesch was published, 220 experienced teachers and other school people from 33 different States of America were asked by Russell what place they actually gave to phonics in their teaching of reading. According to the replies, 15 per cent. gave phonics instruction in lessons regularly scheduled but separate from other reading lessons; 4 per cent. taught phonics in all reading lessons; 39 per cent. in most reading lessons; and a further 39 per cent. as part of a few reading lessons. Only 2 per cent. of the teachers omitted phonics training altogether.

Further questions revealed that whilst in the sample studied, there is no doubt about phonics being used, it is equally clear that they are regarded as only one means of developing word-attack skills, which in turn are viewed as only one goal in the total reading programme.

Similarly, when the present writer recently visited a random sample of 60 Kent primary schools, every school was found to be giving some phonic training in conjunction with other methods, but with different emphasis according to the children's needs and the beliefs of their teachers. In 13 schools children began learning to read by phonic methods, and in six of this group they were only five years old when introduced to them. Moreover, the head teachers of two-thirds of the 60 schools stressed that backward readers in the transition classes and where necessary further up the school, were receiving a great deal of phonic instruction.

There is thus some evidence that phonic skills are being taught in schools at the present time, although prolonged classroom observation in a large and representative sample of schools is necessary before the true position of phonics may be ascertained.

Granted that most teachers use phonic methods today it does not necessarily follow that they should do so. It will be remembered that the value of the alphabetic method for spelling blinded teachers for centuries to its relative ineffectiveness as a method of teaching reading.

The Value of Phonic Methods

Phonic methods have long been the subject of controversy and many investigations have been carried out in an attempt to establish their value.

One of the earliest studies was that of Gill who, in 1912, reported that he had compared the reading ability of three groups of top infants from schools using three different methods, namely, the 'Dale' method of phonic instruction, a phonetic method in which the forms of the letters had been made attractive by ingenious but irrelevant associations, and the sentence method. The first two methods had been used for two years and the third for sixteen months. Reading ability was measured by the time required to read two passages taken from an infant primer but printed without punctuation or spacing between the words, and in one case with the initial letters of nouns printed in capitals. Comparing the average times, the fluency and intelligence of reading of the three groups, Gill concluded that the sentence or whole-word method was of greater practical value than either of the two phonic methods.

Since Gill's report was published at a time when phonic methods were still in the ascendancy his findings were rather disconcerting. However, a year later, in a paper reporting his own experiments, Valentine stated that little significance could be attached to the results of this study since the comparative intelligence of the children was unknown, the general efficiency of the teachers had not been considered, and the form in which the passages were printed favoured the group which had learned to read by sentence wholes. Later investigators have also pointed out that the sentence method group was further favoured by the use of speed as the criterion of reading ability, and by the fact that children in this group were older and presumably more 'ready' for instruction when it was given.

In his own research, Valentine compared two groups of children, aged six-and-a-half to eight years, in their ability to read English written by means of a Greek alphabet after having been taught by a phonic method and a look-and-say approach. He concluded that:

> "There seems to be nothing so inherently difficult, even for little children, in the synthetic work involved by the Phonic method as has been asserted by some of its critics. . . . Children taught by the Phonic method do better than those taught by the 'Look and Say' method both in reading words previously seen and words previously unseen, and even in some words seen by the 'Look and Say' children but not seen by the Phonic children. There seems to be some evidence, however, that for very dull children the 'Look and Say' method is more efficient."

An experiment involving the teaching of Greek letters to children with some experience of the English alphabet may be criticised as artificial but at least Valentine made sure of the initial comparability of his samples. Furthermore, by drawing attention to the importance of experimental design and to the fact that a given method might not be equally effective with all pupils, this investigation made a valuable contribution to further studies as well as providing evidence favourable to phonic methods.

It would be impossible, even if it were desirable, to describe the numerous inquiries provoked by the conflicting results of Gill and Valentine. Mention should be made, however, of a frequently quoted experiment by Winch who, in a period when whole-word methods were gaining ground in the schools, produced further evidence of the value of phonic methods. For this experiment, designed to determine the relative effectiveness of the alphabetic, phonic and look-and-say methods as well as a special phonic method called 'phonoscript', two groups of 32 children between five and six from a school in a poor district of London were chosen. They were equated on the basis of preliminary tests of ability to learn associations between shapes and sounds, and then given twenty-five formal reading lessons. At the end of this period the children were tested individually for their ability to read material consisting of words they had already learnt and a certain number of new words.

Results favourable to phonics and to the more recent system of 'phonoscript' in particular were obtained. Perhaps of even greater interest to teachers, however, was

Winch's final recommendation that, "English children under our conditions of school organisation, may quite profitably begin definitely to learn to read at an average age of 5 years 3 months. . . . In schools where children are well born and of good environment they may begin earlier; in schools attended by very poor children, later." Conclusive evidence of the relationship between socio-economic background and reading ability has since been obtained, but this is one of the first occasions when a research worker pointed to the necessity of adapting teaching methods to the home environment.

The way in which this inquiry was conducted has been praised highly, but it might be said that twenty-five lessons spread over a two-month period were too few to give a reliable comparison of methods, and that, if the lessons had continued, the initial advantage of the phonic methods might have been lost to the more interesting whole-word approach. Further, as in all experiments where one teacher is responsible for all of the teaching, individual differences in the skilful use of a particular method must be taken into account when assessing the results.

In the period following Winch's investigation most of the experimental work on reading methods was carried out in the U.S.A. There, Mosher and Newhall reported in 1930 that there was no essential difference between the effects of a phonic and word approach at least at the end of the second school year. One might question this finding, however, in that accuracy in word recognition was not tested in their investigation and this is the aspect of reading skill to which phonics is supposed to contribute most directly. On the whole, the verdict of American research seems to be in accord with Valentine and Winch in finding that phonic methods are of value, although one authority, Gates, has stated that if it were necessary to choose between phonic and non-phonic training the latter would be selected.

In recent years, the problem has become not the simple black and white one of phonics versus no phonics but rather the more complex one of when, how much, for what purpose, and to whom phonics should be taught. A final verdict on some of these questions has not yet been obtained, but consideration of a few of the more frequently quoted studies may give some idea of the present state of our knowledge about phonic methods. Only reference to researches carried out in English-speaking countries will be made since the effectiveness of phonic methods is obviously determined to a large extent by the nature of the language concerned.

The need for phonic readiness

When experiments on reading methods were first conducted, it was customary to start all children with phonics from the beginning; in current practice phonic methods are not usually introduced formally until a later stage in the infant school course. How far research favours this practice may be seen from the following investigations.*

In 1925, nearly one thousand pupils in eight schools of various types from different sections of Newark, New Jersey, were divided into phonic and non-phonic instructional groups. Their teachers were required to alternate between the methods in an attempt to obviate the effect of differences in teaching efficiency. At the end of five months of training and again at the end of ten months, the children were given a number of reading tests and the results for those pupils who had received uninterrupted instruction were compared. Further tests were also made on 426 pupils who continued in controlled groups through the first half of the second grade.

---

*In evaluating the evidence of research carried out in other countries it is important to be aware of differences in school organisation. For example, only about 50 per cent. of five-year-old children in the U.S.A. are enrolled in school compared with the total child population in this country. American elementary schools usually have either six of eight grades and, where kindergartens are not maintained, children enter the first grade at about the age of six years.

According to the experimenters, Sexton and Herron, the results clearly indicated that the teaching of phonics functions very little or not at all with beginners in reading during the first five months, but begins to be of some value during the second five months, and is of greater value in the second grade. There was also evidence that the teacher is more important than the method, for where results were unusually good in a class taught by a teacher using phonics they were also unusually good when the same teacher taught without phonics. On the other hand, poor results were obtained in both phonic and non-phonic groups taught by the same teacher.

A few years later, Garrison and Heard carried out a three-year sequential study of 111 first-graders, divided into four groups on the basis of intelligence test scores. In four different classrooms with approximately equal school surroundings and quality of teaching, two groups received phonic instruction and two had none.

These investigators came to the conclusion that phonic training should be deferred until the second or third grade (i.e., pupils aged from 7-8 years, approximately). They also reported that whilst instruction in phonics develops independence in word pronunciation, children with no phonic training make smoother and better oral readers in the lower grades. Moreover, bright pupils appeared to be helped more by phonic instruction than those who were dull, and, although children with phonic training had some advantage in learning to spell, the non-phonic first-graders lost less during vacations.

Whilst experiments have shown that it is unprofitable to introduce all beginners to phonics, nevertheless it is reasonable to suppose that children of high intelligence and good cultural background might be expected to start earlier. More recent research on the optimum time for beginning phonic training, therefore, has largely been directed towards ascertaining the relationship between phonic ability and mental maturity. As yet, data on this topic are rather meagre, but such evidence as exists confirms that these two characteristics are highly correlated.

Dolch and Bloomster, for example, administered tests of phonic ability to first and second grade children in a school where phonics instruction had been given but not to an unusual degree. The results showed that, although able children sometimes fail to acquire phonic ability, those with a mental age below seven consistently fail to do so. When the experiment was repeated a year later, substantially the same results were obtained and so it was concluded that a mental age of seven years is necessary for a child to make the best use of phonics.

These findings are confirmed by Schonell and many other investigators, who likewise agree that certain developments and acquisitions are necessary before phonic training should be undertaken. According to Harris the important aspects of phonic readiness in increasing order of difficulty are as follows:

1. The child should be able to hear the difference between words that sound somewhat alike, such as *man* and *men*, or *had* and *hat*.
2. He should be able to detect whether two words begin with the same sound and supply two or three other words that begin with the same sound.
3. He should be sensitive to rhymes, able to pick out words that rhyme, and supply words to rhyme with a given word.
4. He should be able to hear similarities and differences in word endings and middle vowels of words.
5. He should be able to listen to the pronunciation of a word sound by sound, and fuse or blend the sounds mentally so as to be able to recognise the word intended.

Laboratory experiments have demonstrated that auditory discriminations, such as those listed above, are usually more difficult and therefore acquired at a later stage than those involving visual perception. Further, the visual perception of whole words has been shown to precede the ability to differentiate between the individual letters which form them. This additional evidence of the development of perceptual skills reinforces the conclusion reached by research in the classroom that *formal* phonic training for a considerable proportion of very young children is largely a waste of time.

The nature and amount of phonics instruction which should be given

Research studies have established the need for phonic readiness but it is nowhere suggested that teachers should not provide for its development in the reception class, or even earlier in the nursery school. Some researchers, including Dolch, Gray, Harris and Watts, have stated that at quite an early age *informal* training in the recognition of similarities and differences in word sounds by means of rhymes, etc., is profitable. It would seem that how phonic training is given is just as important as the timing of its introduction.

In the experiments previously cited Gates, Sexton and Herron concluded that it is preferable to teach sounds in words and not to drill them in isolation. Other investigators agree with this finding but seem to disagree about the amount of phonic instruction which should be given, and whether it should be formal or informal.

In 1938, Gates and Russell reported the results of a study in which 354 first-graders " fairly representative of the population of New York City," were first equated for mental age and then assembled into three groups. Group D " received the smallest amount of phonics or word analysis." Group E was given " moderate amounts of informal, newer type word analysis, comparisons, etc." Group F " had substantial or large amounts of conventional phonic drill."

At the end of the training period Group E had the highest scores in all four tests of word recognition and comprehension, Group D exceeded Group F slightly in two tests and was equal in the rest. It was concluded that a programme including little or no phonic analysis in the first year is not as good as an informal programme but is at least as good as one containing large amounts of formal phonic work.

Tate, Herbert and Zeman also found that the incidental teaching of phonics in connexion with children's needs in working out words encountered in their reading was superior to the teaching of isolated phonics or no phonics instruction.

When Agnew considered the effects of large amounts of phonic training on grade one children, he too found a slightly detrimental effect. In grade two the effect seemed to be one of slight reading improvement and in grade three of more accurate pronunciation. Since he was unable to establish a critical grade in which phonic experience is particularly effective, a further experiment was carried out. This time he reported that third-grade children who had experienced greater amounts of phonic instruction excelled, by statistically reliable differences, in tests of word pronunciation, vocabulary, accuracy of oral reading and eye-voice span, and yet there was comparatively little difference between them and pupils receiving a small amount of phonic instruction in the silent reading abilities measured.

The kind of reading skill developed by systematic phonic instruction

Important as it is to be able to recognise words in order to read them, it might be argued that comprehension is the essence of reading and that a lopsided development of word recognition skills has no merit, *per se*.

This argument is supported by the work of Tate who concluded that: "The use of as many as thirty minutes daily for special phonics instruction and drills leads to an unbalanced development of the abilities to comprehend words, to understand sentences and to grasp the meaning of paragraphs."

More recently, a similar conclusion was reached by McDowell in a study of fourth-graders of comparable intelligence and socio-economic background from two groups of Pittsburgh schools. He reported that in the group of schools where phonic training was introduced gradually, developed through the analysis of meaningful material and treated as a subsidiary word-attack skill, pupils were able to read faster, understand words and paragraphs, use an index, and generally read better than those in schools employing a conventional phonics programme. The phonic group was clearly superior in alphabetizing and spelling but, according to the experimenter, this superiority had been achieved at the expense of a more rounded and harmonious development of reading skills.

These researches and the work of other investigators, including that of Garrison and Heard and Gates mentioned earlier, would seem to conflict with Agnew's finding that improved word recognition through formal phonics training does not result in a lowering of reading comprehension ability.

There is, however, some evidence to support Agnew's conclusions in the recent investigation conducted in Kent primary schools by the National Foundation (Morris). Here, children aged seven to eleven who had attended schools which concentrated on systematic phonic instruction at the infant stage, were markedly superior in tests of reading *comprehension* to those who had been pupils at schools where phonic training had been given incidentally and without emphasis. These results were obtained under ' normal ' school conditions (i.e. without recourse to ' artificial ' experiment), and after allowance had been made for the children's non-verbal ability, and three school characteristics, namely, urban/rural location, size of school, and the socio-economic status of the school catchment area.

As Daniels and Diack suggest, it seems likely that much of the controversy arises because different varieties of the phonic method have been used. It is also probable that sufficient attention has not been paid to differences between the immediate and after-effects of these methods, for Tiffin and McKinnis discovered that there is a high positive correlation between reading comprehension ability and phonic ability when children are in grades five to eight. They therefore concluded that a reading programme "which does not, by direct or indirect instruction, yield a mastery of the principles of phonics is not accomplishing its full purpose."

The type of pupil who benefits most from phonic teaching

Phonic ability is not necessarily the result of satisfactory school methods for, according to Schonell and others, intelligent children need comparatively little phonic teaching since they are able to understand and blend word sounds for themselves. It is children of poor intelligence, physique, emotional stability and socio-economic status for whom careful attention to methods is important.

It will be remembered that as early as 1913 Valentine showed that the dull child might make greater progress with look-and-say rather than phonic methods in the early stages. Ten years later Currier likewise pointed out that whilst phonic drills are valuable they are not essential to every child, and that at all times they should be employed with discretion and adapted to the needs of individual children or special groups.

Experiments designed to find out more about phonic readiness have indicated that formal phonic training should be delayed until a certain maturity has been achieved. This means that some children at seven plus, or even older, are only just ready to receive systematic phonic instruction, but that if this is not given, satisfactory reading progress may not be made. Some investigators have therefore turned their attention to the problem of phonic training in relation to children beyond the infant stage. In this connexion, Daniels and Diack compared the effects of teaching reading to first-year junior non-readers by the modified phonic method used in the *Royal Road Readers*, and a 'mixed-methods' approach beginning with look-and-say methods supplemented at some stage by phonic teaching. After one year the children were given six specially designed tests and it was found that those taught by the phonic word-method were markedly superior in word recognition and word attack to pupils in the mixed-methods group.

The value of phonic teaching with backward readers in second year junior B and C classes has also been shown in an investigation by Conduct and Ward. Besides their ordinary reading of books the children were given apparatus, games, and tests of phonic sounds, and were also taught to recognise sounds in words and how to attack new words. After a ten-month period of teaching, which included individual, group and class activities, the reading ages of the 44 pupils were increased by 0.8-3.0 years, the average increase being 1.5 years. Furthermore, all achieved some degree of literacy although some of the children were almost non-readers when the experiment began and several of them were sub-normal in intelligence.

Not always has the effectiveness of phonic methods with backward readers been so clearly demonstrated. In fact, after experimenting with ten to eleven-year-olds drawn from their normal classroom environment, Burt and Lewis reported that, on the whole, the best method of teaching dull and backward readers is one which is predominantly visual. They pointed out, however, that the differences found in their experiment between whole-word, alphabetic, phonic, kinaesthetic and 'mixed' methods, even when significant, are still not very large, and that there is an urgent need for more intensive research into the problem.

It is possible, as Conduct and Ward suggest, that research into the effectiveness of different remedial measures has been directed too much to the individual and too little to the class. Individual treatment is known to give positive results whatever the methods used. What inexperienced teachers need to know is which methods are most effective in ordinary classroom conditions where it is often difficult to give sufficient attention to the needs of individual children. As will have been observed, some progress has been made towards answering this complex question in the limited field of phonics.

Summary

We now know that phonic instruction is of value and should be included in the reading programme. To be successful, however, it must be regarded as only one method of developing word skills and should not be introduced formally until children are ready for it. Bright pupils usually will have acquired sufficient vocabulary and reached sufficient maturity to be able to start phonic analysis in the reception class, but for those who are dull or handicapped in any other way, formal phonic training can profitably be postponed until a later stage in the school course—even into the first or second junior school year. There is still some doubt about the amount of phonic training which should be given but convincing evidence has been obtained that sounds should be taught in words and not drilled in isolation. Phonic methods undoubtedly improve word recognition but it is not certain whether this is gained at the expense of comprehension and interest in reading; the high positive correlation between phonic ability and reading comprehension ability in older children would suggest that this might not be so. Finally, the effectiveness of phonic methods with backward readers beyond the age of seven is indicated by the few experiments which have been carried out in conditions approximating to those of a normal class.

Bibliography

AGNEW, D. C., "The Effects of Varied Amounts of Phonetic Training on Primary Reading." *Duke University Research Studies in Education*, No. 5. Durham, N.C.: Duke University Press, 1939.

BURT, C., and LEWIS, R. B., "Teaching Backward Readers." *Brit. J. Ed. Psych.*, Vol. XVI, Part III, November, 1946.

CONDUCT, G. N., and WARD, H., "Phonic Reading—A New Approach." *Educational Review*, Vol. VII, No. 3, June, 1955.

CURRIER, L. B., "Phonics or No Phonics." *The Elementary School Journal*, Vol. XXIII, February, 1923.

DANIELS, J. C., and DIACK, H., *The Royal Road Readers*. London: Chatto & Windus, 1954.

DANIELS, J. C., and DIACK, H., *Progress in Reading*. University of Nottingham Institute of Education, 1956.

DOLCH, E. W., *Problems in Reading*. Champaign, Illinois: Garrard Press, 1948.

DOLCH, E. W., and BLOOMSTER, M., "Phonic Readiness." *The Elementary School Journal*, Vol. XXXVIII, November, 1937.

FERNALD, G. M., and KELLER, H. B., " The Effect of Kinaesthetic Factors in the Development of Word Recognition in the Case of Non-Readers." *J. Educ. Res.*, Vol. IV, 1921.

FLESCH, R., *Why Johnny Can't Read.* New York: Harper, 1955.

GARRISON, S. C., and HEARD, N. T., " An Experimental Study of the Value of Phonetics." *Peabody Journal of Education*, Vol. IX, July, 1931.

GATES, A. I., " Studies of Phonetic Training in Beginning Reading." *J. Ed. Psych.*, Vol. XVIII, April, 1927.

GATES, A. I., and RUSSELL, D. H., " Types of Material, Vocabulary Burden, Word Analysis and Other Factors in Beginning Reading." *The Elementary School Journal*, Vol. XXXIX, 1938.

GILL, E. J., " Methods of Teaching Reading: A Comparison of Results." *J. Exp. Pedagogy*, Vol. I, No. 3, March, 1912.

GRAY, W. S., " The Nature and Organisation of Basic Instruction in Reading." *The Teaching of Reading: A Second Report*, 36th Yearbook of the U.S.A. National Society for the Study of Education, Part I. Bloomington, Illinois: Public School Publishing Co., 1937.

GRAY, W. S., " Reading." *Encyclopedia of Educational Research*, 1950.

HARRIS, A. J., *How to Increase Reading Ability.* New York: Longmans Green, 1947.

McDOWELL, J. B., " A Report on the Phonetic Method of Teaching Children to Read." *The Catholic Educational Review*, Vol. LI, October, 1953.

MORRIS, J. M., *Reading in the Primary School.* Publication No. 12 of the National Foundation for Educational Research. London: Newnes (*in the press*).

MOSHER, R. M., and NEWHALL, S. M., " Phonic Versus Look-and-Say Training in Beginning Reading." *J. Ed. Psych.*, Vol. XXI, October, 1930.

RUSSELL, D. H., " Teachers' Views on Phonics." *Elementary English*, Vol. XXXII, October, 1955.

SCHONELL, F. J., *Backwardness in the Basic Subjects.* Edinburgh: Oliver & Boyd, 1942.

SCHONELL, F. J., *The Psychology and Teaching of Reading.* Edinburgh: Oliver & Boyd, 1947.

SEXTON, E. K., and HERRON, J. S., " The Newark Phonics Experiment." *The Elementary School Journal*, Vol. XXVIII, May, 1928.

TATE, H., " The Influence of Phonics on Silent Reading in Grade I." *The Elementary School Journal*, Vol. XXXVII, June, 1937.

TATE, H., HERBERT, T., and ZEMAN, J., " Non-Phonic Primary Reading." *The Elementary School Journal*, Vol. XL, March, 1940.

TIFFIN, J., and McKINNIS, M., " Phonic Ability: Its Measurement and Relation to Reading Ability." *School and Society*, Vol. LI, February, 1940.

TRAXLER, A. E., *Ten Years of Research in Reading.* Educational Records Bulletin, No. 32. New York: Educational Records Bureau, 1941.

TRAXLER, A. E., and Townsend, A., *Another Five Years of Research in Reading.* Educational Records Bulletin, No. 46. New York: Educational Records Bureau, 1946.

TRAXLER, A. E., and TOWNSEND, A., *Eight More Years of Research in Reading.* Educational Records Bulletin, No. 64. New York: Educational Records Bureau, 1955.

VALENTINE, C. W., " Experiments on the Methods of Teaching Reading." *J. Exp. Pedagogy*, Vol. II, No. 2, June, 1913.

VERNON, M. D., *Backwardness in Reading.* Cambridge University Press, 1957.

WATTS, A. F., *The Language and Mental Development of Children.* London: Harrap, 1944.

WINCH, W. H., " Teaching Beginners to Read in England: Its Methods, Results and Psychological Bases." *Journal of Educational Research Monographs*, No. 8. Bloomington, Illinois: Public School Publishing Company, 1925.

A COMPARATIVE STUDY OF BASIC ATTAINMENTS

The Results of Surveys of the Attainments in Reading and Arithmetic of Eleven-year-old Children in England and Wales; Queensland, Australia; and Central California, U.S.A.

by D. A. Pidgeon

I. Introduction

IN May, 1955, as part of a national survey involving three age groups, a representative sample of some 3,000 children, aged eleven plus, attending primary schools in England and Wales, were given specially prepared tests of reading, arithmetic and non-verbal ability. After reading an account of the proposed survey in the Foundation's Bulletin No. 5, the Principal Research and Guidance Officer for the State of Queensland, Australia, sought permission to apply the same tests to a comparable sample of eleven-year-old children in Queensland. The request was granted and in November, 1955, the survey was repeated on a representative sample of children from that State. In the summer of 1956, following a visit to the Foundation by Professor G. T. Buswell of the University of California (Berkeley), U.S.A., arrangements were made for one of the arithmetic tests used in the Foundation's survey to be repeated in Central California. Accordingly, in January, 1957, this arithmetic test was given to a representative sample of eleven-year-old Californian children.

In the following pages an account is given of a comparision of the results obtained from these three surveys.

II. The Samples Tested

(*a*) England and Wales

The national survey was confined to grant-aided schools maintained or assisted by local education authorities in England and Wales, and in this particular age group was restricted to children attending junior and all-age schools who had reached their tenth birthday but not their eleventh by August 31st, 1954, that is, who were born between September 1st, 1943, and August 31st, 1944 inclusive. Since the testing was carried out in May, 1955, the effective age range of the children was from 10 years 9 months to 11 years 8 months. A random sample of schools was drawn and all children in the age group in each selected school were tested. The selection of schools was made directly from the lists kept at the Ministry of Education of all schools in England and Wales attended by the children concerned. In order to ensure the adequate representation of schools of different types and in urban and rural areas, random sampling was carried out separately in each of these strata. The sampling method was thus systematic but unbiassed and was considered to be effectively random. In all, some 3,180 children attending 91 schools administered by 63 different local education authorities were tested.[1]

---

[1] Full details of the sampling procedure employed will be published later.

(*b*) Queensland, Australia

Surveys of children's abilities and attainments had previously been carried out in Queensland and therefore the technique of obtaining effectively random samples was not new to them. In this inquiry, as in England and Wales, the schools selected for the survey were a stratified random sample. The testing was confined to children born between March 1st, 1944, and February 28th, 1945, and to equate conditions with regard to the school year in this country the testing was carried out in November, 1955. Thus, the mean age of the sample was 11 years 2½ months and the fact that the tests were given towards the end of the school year made it virtually identical with that used in England and Wales. The schools in Queensland containing children in the appropriate age group were divided into the five groups according to size that are used officially for purposes of classification. The groups are: Class I, average attendance over 400; Class II, 150-400; Class III, 41-150; Class IV, 21-40; Class V, 20 and under. The number of schools selected in each class was fixed at 1/17th of the number of schools in that class in 1954. As this would have led to the selection of only six schools of more than 400 pupils, it was decided to select three times as many Class I schools but to test only one-third of the children in each. The grade system of placement is used in Queensland schools and although approximately two-thirds of the children in the defined age group were in Grade V the others were found as low as Grade II (two pupils) and as high as Grade VII (one pupil). In all some 940 children attending 88 schools were tested.

(*c*) California, U.S.A.

The sampling procedure in California was very similar to that employed in England and Wales. Again, although the grade system of placement is used in America, the testing was confined to children of the same age range at the date of testing as those in this country and in Queensland. Thus, the sample was drawn from all pupils who were born between May 1st, 1945, and April 30th, 1946, inclusive, and the testing was carried out in January, 1957. The 773 urban and 661 rural schools in the 21 counties in Central California were numbered serially, and after stratifying according to urban or rural locality an effectively random sample of schools was selected to give a sample of children approximately equal in number to that obtained in England and Wales. Invitations were issued to the selected schools to participate in the experiment and only three of the school systems refused. One or two other minor exceptions to the sampling procedure were necessary, but it is unlikely that they exerted any appreciable effect upon the results obtained. In all, 3,179 pupils from 70 schools were tested.

III. The Tests

Four tests were administered in the Foundation's national survey to children in the 11 + age group. These were:

1. *A test of non-verbal ability*, consisting of 64 items of two main types. The first type consisted of a large square with the lower right-hand corner missing, and several smaller squares from which the appropriate one had to be selected to complete the large square. The second consisted of a sequence of three or four drawings with the next one missing, and a group of five more from which the appropriate one had to be chosen to continue

the sequence. The items were arranged four to a page, but they were not grouped in separate series. A practice test of eight items was administered with detailed oral instructions before the test proper. The reliability of the test was found to be 0.943.*

2. *A reading comprehension test*, consisting of 60 incomplete sentences. The appropriate word necessary to complete the sentence had to be chosen from five alternatives. The items, presented in order of difficulty, ranged from sentences suitable for bright seven-year-olds to those found difficult by children of fourteen. The test was scored by counting all correct items up to the point at which a gap of seven consecutive wrong answers occurred. Any items correct beyond this point were assumed to be due to chance and ignored. The reliability of this test was 0.958.*

3. *A mechanical arithmetic test*, consisting of twenty-five items ranging from a two row addition sum with hundreds, tens and units, and a carrying figure in the units column, to a simple multiplication of decimals. Ten sums dealt with number facts only, the most difficult being a long multiplication and a long division without remainder. A further nine sums dealt with conversion units, involving money, time, length, weight and capacity. The last six items consisted of four elementary fraction sums, one in each basic process, and two simple decimal sums. The test was administered without time limit. The reliability of this test was 0.922.*

4. *A mechanical and problem arithmetic test* consisting of 100 items administered with a time limit of 50 minutes. The test was in two sections, separately timed. Section 1 consisted of 40 mechanical sums involving the four rules and exercises in money, weight, time, length and capacity. Section 2 consisted of 60 simple one-step problems covering the same ground with the addition of one or two involving elementary fractions and area. The reliability of this test was .982.*

These four tests were also given in the same order and with the same instructions in Queensland, Australia.

In California only the second 100 item arithmetic test was administered. Certain changes were necessary in adapting the test for American usage. (*a*) Seventy of the 100 items were considered suitable for use in American schools with only minor verbal changes. (*b*) Twenty-six items involved money terms. These were changed to dollars and cents, with due regard to the size of the numbers and the arithmetical processes involved. (*c*) The remaining four items involved weight measured in "stones"—a unit not used in America. In tabulating the results, both in California and in this country, the scores obtained in these three groups of items were treated separately. For this present report an analysis was carried out on the 70 comparable items only, of which 31 were mechanical sums in Section 1, and 39 were problems from Section 2. On the data from England and Wales, the reliability was .974 for the total of 70 items, and .949 for Section I and .955 for Section 2 separately. On the Californian data the reliability was .89 for the total of 70 items and .68 for Section 1 and .87 for Section 2 separately.*

---

*By the reliability of a test is meant the degree to which scores derived from it are likely to be consistent. Measures of reliability are usually expressed in a scale from 0 to 1.0. The measures for these tests were obtained by Kuder-Richardson formula 20.

IV. Administration of the tests

(*a*) England and Wales

In the Foundation's national survey the testing was carried out, either by the head teacher of the school concerned, or by a psychologist or other trained official of the local education authority concerned. In the case of the former, responsibility for the administration of the tests rested nevertheless with an official of the local education authority. A detailed instruction booklet with clear and simple directions for giving the tests had been prepared. The tests were administered on the mornings only of three consecutive days; the non-verbal test was given on the first day, the reading test, followed after a short break by the untimed arithmetic test, was given on the second day, and the timed arithmetic test was given on the third day.

In order to ensure that the testing was properly carried out each local authority concerned was asked to appoint an official to be responsible for the conduct of the survey locally. These officials were briefed by members of the Foundation at a series of regional meetings held for the purpose at eight centres scattered over the country. Reports on the conduct of the survey locally were asked for and a study of these indicated that in most cases a great deal of trouble had been taken to ensure that the children worked in circumstances that enabled them to do their best. Although in particular instances, especially in small rural schools, difficulties of accommodation were experienced, it appeared that these were overcome, and reliance could be placed upon all scores obtained. Completed tests were returned to the Foundation for marking and checking.

(*b*) Queensland, Australia

In Queensland the order of giving the tests and the instructions for administration were identical with those in England and Wales, and considerable care was taken to ensure that conditions were satisfactory. The tests were administered by teachers, but the marking and checking was carried out by the staff of the Research and Guidance Branch of the Department of Public Instruction.

(*c*) California, U.S.A.

In California all the tests were administered by Professor Buswell or his research assistant, George Yonge. Generally, one or more teachers or supervisors were also in the room during the testing and the physical conditions were reported to be good. The instructions for the test were given orally and were uniform with those used in England. In some of the schools the test was given in the regular classrooms, pupils outside the age range being sent to other parts of the building, while in the larger schools the pupils to be tested were assembled in a multi-purpose room using cafeteria or library tables. All the tests were scored by clerical assistants in the research office at Berkeley University. Their work was checked systematically to ensure correct scoring.

V. Results and Comments

(*a*) Comparison between England and Wales and Queensland

As only one test was given in California, the results of the full survey common to both this country and Queensland will be given first.

The basic sample of children considered in each country was that consisting of those children who completed all four tests. In England and Wales, the data obtained from all children who completed each test were, however, used to standardise the tests. The method used for the standardisation was an adaptation of that devised by Lawley[1] in which the standardised scores are so arranged that they have a mean of 100 and a standard deviation of 15.[2] Using the conversion tables prepared, the mean standardised scores of the sample of children taking all tests were thus calculated for both England and Wales and Queensland. Table 1 gives these means (standard errors in brackets), together with the respective standard deviations.

Table 1

COMPARISON OF MEAN STANDARDISED SCORES FOR FOUR TESTS GIVEN IN QUEENSLAND, AUSTRALIA AND IN ENGLAND AND WALES

	N/V Test		Reading Test		Untimed Arithmetic Test 1		Timed Arithmetic Test 2	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Queensland, Australia	94.7 (0.66)	12.5	95.3 (0.65)	12.7	100.3 (0.54)	10.6	99.7 (0.57)	10.4
England and Wales	100.48 (0.80)	14.50	100.03 (0.69)	14.51	100.32 (0.80)	14.88	100.39 (0.77)	14.75

The slight discrepancies between the figures given for England and Wales in Table 1 and the rounded values of 100 and 15 for the means and standard deviations respectively are due in the main to the omission of partial absentees. The mean standardised scores of those children who were present for one test, but absent for one or more of the others, were in fact calculated and found to be some 4 or 5 points below the value of 100.

The main fact emerging from Table 1 is the significantly lower mean scores obtained by the Queensland children on both the non-verbal and reading tests. The differences between the means for both arithmetic tests are not significant. The appreciably smaller standard deviations are also worthy of note.

In any discussion of these findings two comments received from Queensland are relevant. In the first place it was pointed out that this sample of Australian children were completely lacking in test sophistication and that this fact might affect the scores on the non-verbal test more than those on the other tests. Secondly, previous Australian comparisons had shown Queensland pupils to be well ahead of those in the other States on standardised arithmetic tests (of a type similar to the untimed Arithmetic Test 1) and average on tests of silent reading and non-verbal intelligence.

It may well be true that the English sample of children had enjoyed a wider experience of tests in general, although from the information gleaned from the survey reports sent in by many local authorities it is unlikely that many of the children tested had previously been given non-verbal tests or tests of reading similar to the one employed in the survey.

---

[1] *A Method of Standardising Group Tests* by D. N. Lawley, B.J.P. (Statistical Section), Vol. III, Part II, June 1950.

[2] The 'standard deviation' is a measure of the spread of the scores—i.e. the extent to which they tend to deviate from the mean or average

The non-verbal test

It is possible, however, that the children in England and Wales were more familiar with the type of perceptual material employed in the non-verbal test by reason of the increasing number of puzzle toys now available for children of primary school age. In any case, it is important to remember that this test is measuring only one aspect of intelligent behaviour, and the different cultural environments of the children in the two countries may well be responsible for the difference in the scores obtained. It cannot be assumed, therefore, that the lower scores of Queensland children in any way reflect a lower innate intelligence.

It will be observed in Table 1 that the standard deviations of the standardised scores for the non-verbal and the other three tests are somewhat smaller in the Queensland data than they are in England and Wales. A more detailed comparison of the distribution of test scores in the two countries was made and Table 2 gives the proportions found in three groups of standardised score. The expected proportions from a normal distribution are also given.

Table 2

PROPORTIONS IN QUEENSLAND AND IN ENGLAND AND WALES OCCURING IN THREE GROUPS OF STANDARDISED SCORE, ON FOUR TESTS

Range of Standardised Scores	Expected Proportion from Normal Distribution	Non-Verbal Test		Reading Test		Arithmetic Test 1		Arithmetic Test 2	
		Q.	E.	Q.	E.	Q.	E.	Q.	E.
115 & above	15.87	5.73	17.20	5.73	14.89	8.38	17.14	6.04	16.54
85–114	66.52	76.88	68.66	75.40	69.63	85.47	68.47	85.47	68.17
84 & below	17.62	17.40	14.13	18.87	15.48	6.15	14.40	8.48	15.28

It is clear from Table 2 that the lower mean score on the non-verbal test for Queensland children was due, in the main, to the relatively smaller proportion of high scorers.

Reading Test

So far as the reading test is concerned the possibility that some of the words used were inappropriate for Australian children cannot be overlooked. It was ascertained from the Principal Research and Guidance Officer, Queensland, however, that although an occasional word or concept might be more fitting for English than Australian children, he would hesitate to assert this. An item analysis was carried out on a sub-sample of scripts from each of the two countries, and an inspection of the results showed that in only one item was there a serious discrepancy. This item, answered correctly by approximately fifty per cent. of English children and by only twenty-five per cent. of Queensland children, was concerned with the meaning of the word "conjuror" which presumably is not so familiar in Australia.

The proportions of high and low scorers on this test, given in Table 2, are interesting. Not only is there a much higher percentage of English and Welsh children who do well on this test, but there is a slightly smaller percentage doing badly. It may be safe, therefore, to conclude that in-so-far as this test measured reading comprehension, the children from this country were superior in this respect to those from Queensland.

Arithmetic Tests

The values for the mean scores given in Table 1 on both arithmetic tests would suggest that there is little to choose between the performances of children from the two samples, although the low standard deviations of the Queensland children call for attention. The appropriate figures in Table 2 emphasise the difference between the distribution of scores obtained in the two countries. There are more high scorers on both tests among the English and Welsh children, but there are also more low scorers. In seeking an explanation of these findings it would seem appropriate to examine the different curricula followed by the two groups of children in their respective primary schools. Before doing so, however, it is necessary to look at another factor that might have influenced the results.

If it can be assumed that the sample children from this country were relatively unsophisticated in non-verbal and reading comprehension tests, a similar assumption could not be made with regard to the timed Arithmetic Test 2. This test was similar to that employed by many local education authorities as part of their 11 + allocation procedures, and since nearly all the English children had undergone a selection examination some three months previously, this type of test could not be regarded as unfamiliar. Not all authorities, however, employed standardised arithmetic tests in their allocation procedures, but from the information available at the Foundation it was possible to divide the sample schools into two groups: Group (*a*)—those in authorities that regularly used a standardised arithmetic test, and Group (*b*)—those in authorities that did not. The mean standardised scores, on all four tests used in the survey, of the children attending schools in these two groups are given in Table 3.

Table 3

MEAN STANDARDISED SCORES ON FOUR TESTS FOR:

(*a*) Children in L.E.As. that used standardised arithmetic tests.
(*b*) Children in L.E.As. that did not use standardised arithmetic tests.

	Non-Verbal Test	Reading Test	Untimed Arithmetic Test 1	Timed Arithmetic Test 2
Group (*a*)	101.03	100.18	100.67	101.14
Group (*b*)	99.85	99.86	99.93	99.52
Difference	1.18	0.32	0.74	1.62

None of the differences shown in Table 3 is significant, although, as might have been expected, the largest difference is found with the timed Arithmetic Test 2.

The influence of test sophistication may, however, be seen in another way. Since Arithmetic Test 2 is timed and the rate of working expected is rather high, it is possible that English children as a whole may have been able to tackle more of the test in the time allowed than their counterparts in Queensland. This point was investigated by comparing the average facility values[1] for consecutive groups of ten items obtained from item analyses carried out on sub-samples of the completed scripts. On the English and Welsh sample, a one-in-seven random sample of the original scripts was used for this purpose, and in Queensland the indices of difficulty were obtained from a sample of 150 pupils whose scores on the non-verbal test were representative of those of all pupils.[2] Table 4 gives the average facility value for successive groups of 10 items.

Table 4

MEAN FACILITY VALUE FOR SUCCESSIVE GROUPS OF 10 ITEMS IN ARITHMETIC TEST 2, FOR QUEENSLAND AND ENGLAND AND WALES

ITEMS	MEAN FACILITY VALUES	
	Queensland	England & Wales
Section 1. Items 1–10 ..	76.2%	69.2%
,, 11–20 ..	64.0%	52.3%
,, 21–30 ..	48.6%	38.6%
,, 31–40 ..	16.0%	15.9%
Average for Section 1 ..	51.2%	44.0%
Section 2. Items 1–10 ..	67.6%	69.6%
,, 11–20 ..	42.2%	44.9%
,, 21–30 ..	46.2%	42.6%
,, 31–40 ..	37.6%	37.8%
,, 41–50 ..	15.6%	23.1%
,, 51–60 ..	4.6%	13.3%
Average for Section 2 ..	35.6%	38.6%
Test Average	41.9%	40.7%

There are a number of comments to be made on Table 4. In the first place the mean raw score on this test was 40.5 in both countries. It is clear, therefore, that the Australian sub-sample used in this analysis was slightly superior to the total sample of Queensland children. However, this does not affect the other comments to be made, namely, that the Australian children did considerably better than their English and Welsh counterparts on Section 1—the mechanical section—of the test, were only about equal to them on Section 2—the problem section—and failed, on the average, to get as many items completed towards the end of the test. The assumption, necessary for this last comment, appears to be justified, since there are no specifically new types of item introduced towards the end of the test and the actual increase in difficulty is only slight. It appears, therefore, that, although the mean standardised scores for the two countries on this arithmetic test did not differ significantly, this was due (*a*) to the superior performance of the Australian children on mechanical sums, and (*b*) to the English and Welsh children being able to get more of the problem sums right towards the end of the test.

---

[1] The 'facility value' of an item is denoted by the percentage of children in a sample who answer the item correctly.

[2] As would be expected in this case, the distribution of scores on Arithmetic Test 2 showed a slight regression to the mean, but it was not thought that this would affect the indices to any great extent.

However, the question of rate of working does not enter into the results for Arithmetic Test 1, since it was administered without time limits. Also, since this test consisted of mechanical sums only, a better all-round performance on the test might have been expected from Queensland children. But this is not so, and in order to discover the reason why, it is necessary to examine in greater detail the respective performances of the children from the two countries on specific types of item. Arithmetic Test 1 was, in fact, included in the survey for this very purpose. It is not possible in giving the results of the item analysis to quote the actual sums used in the test, since it was constructed specially for the survey and it is hoped to repeat it on a later occasion. Neither is it considered necessary to detail all 25 items. Table 5, however, gives the comparative proportions answering correctly for 10 selected items. The Australian data are based on the sub-sample of 150 children described previously; the values for the English and Welsh children were derived from all the 3,185 children taking the test.

Table 5

PERCENTAGES OF CHILDREN IN QUEENSLAND AND IN ENGLAND AND WALES ANSWERING CORRECTLY SELECTED ITEMS FROM ARITHMETIC TEST 1

ITEM No.	TYPE OF ITEM	FACILITY VALUES	
		Queensland %	England %
5	Simple add. 3 row h.t.u. 2 carrying figs.	96	92
6	Simple sub. h.t.u. borrowing in 1 col. and zero	83	79
7	Simple mult. h.t.u. by u.	88	77*
8	Short div. th.h.t.u. by t.u. no remainder	89	74*
13	3 row add. hr. min. sec.	66	59
14	Sub. yd. ft. in. borrowing in 1 col.	68	68
15	Simple mult. gal. qt. pt.	55	47
16	Short div. tons cwt. lb. no remainder	61	36*
21	Sub. simple fractions	21	36*
24	Add. of dec. 3 simple numbers presented horizontally	23	43*

Not all the differences shown in Table 5 are significant, but it seems apparent that the Australian children are generally better on all items except those involving fractions and decimals. It will be recalled that there are three main types of item in this test: (*a*) those dealing with simple numbers, involving the 4 rules, including one long multiplication and one long division (10 items); (*b*) those dealing with conversion units, including money, time, length, weight and capacity (9 items); (*c*) those dealing with simple fractions (4) and decimals (2) (6 items). The mean facility values for these three types of item are given in Table 6.

Table 6

COMPARISON OF THE MEAN FACILITY VALUES FOR THREE TYPES OF ITEM BETWEEN CHILDREN IN QUEENSLAND AND IN ENGLAND AND WALES

No. of Items	Type of Item	Mean Facility Values	
		Queensland	England & Wales
10	Simple number: 4 rules	87.1%	79.0%
9	Conversion units. Money, time, length, weight and capacity	65.4%	58.1%
6	Simple fractions and decimals	17.7%	34.3%

The pattern of responses to items of mechanical computation is now quite clear. The children from Queensland give a superior performance to English and Welsh children on both arithmetic tests in all mechanical items except those involving simple fractions and decimals. The latter are not introduced in the schools in Queensland until the Sixth Grade and thus only about eighteen per cent. of the children in the sample had had any formal instruction in sums of this type. Had questions of this kind not been included, therefore, the Australian children would have obtained a higher mean score on Arithmetic Test 1 than the children from this country.

This finding is of interest—not because the English and Welsh children do better than Queensland children on items involving fractions and decimals, but because their success on this type of item is as high as it is. The fact that about one-third of the children from England and Wales were able to answer the fraction and decimal sums correctly is surprising, since this type of item does not appear in the mechanical section of the standardised tests of arithmetic used by the majority of authorities in their 11 + allocation procedures (cf. Arithmetic Test 2). It might have been expected, by reason of the alleged " backwash " effect of such tests, that exercises of this kind would not be taught in English primary schools, but the evidence from the survey is clearly to the contrary.

Teaching mechanical computation is not, of course, the sole constituent of the arithmetic curriculum in primary schools and the necessity for pupils to be able to answer written problems is equally important. Reference to the mean facility values (given in Table 4) for the first forty items of the second (problem) section of Arithmetic Test 2 would indicate little difference in the performances of children from Queensland and England and Wales. The variation in the type of arithmetical problem given in this test is, however, quite large and the grouping adopted for Table 4 does not give any indication of possible differences that may occur between different types of problem. It would be desirable, therefore, to compare the average facility values for items of a similar kind, but in practice this is very difficult to achieve since the majority of problems can be classified under a number of different headings.

An inspection of the separate facility values, however, revealed interesting differences, and to illustrate these the first forty items of the second section of the test were divided into two main groups. Three items were omitted from this analysis since they involved fractions and, for the reasons stated above, a comparison involving those items would clearly have been inappropriate. The last twenty items were also omitted since it was clear that in these the Queensland children were handicapped by the imposition of strict time limits. The two groups, the mean facility values of which are given in Table 7, were as follows: Group (*a*) Eighteen simple written exercises involving the four rules and conversion units, e.g., item 2, " John has 7 sweets, and I have three times as many. How many have I? "[1] Item 6, " Write in figures the number eleven thousand two hundred and seven." Item 38, " Take 144 pence from 47 shillings and give your answer in £ s. d." Group (*b*) comprised nineteen more complex verbal problems, e.g., Item 4, " I bought a three shilling book of stamps and four 2½d. stamps. What was the total cost? " Item 16, " A man cycles at 12 miles an hour. How long will he take to travel 84 miles at this rate? " Item 39, " I can pay 8d. for a return ticket every day or buy a weekly ticket costing 4/-. How much shall I *lose* each week by buying a weekly ticket if I go to work only 5 days a week?"

Table 7

COMPARISON BETWEEN THE MEAN FACILITY VALUES FOR TWO GROUPS OF ITEMS FROM ARITHMETIC TEST 2 FOR QUEENSLAND AND ENGLAND AND WALES

	No. of Items	Type of Item	Mean Facility Values	
			Queensland	England & Wales
Group (*a*)	18	Simple written exercises in 4 rules and conversion units	58.9%	51.1%
Group (*b*)	19	More complex verbal problems ..	41.1%	46.1%

Of the eighteen items in Group (*a*) fourteen were answered correctly by more of the Australian children, and four by more of the children from England and Wales. Of the nineteen items from Group (*b*) the English and Welsh children did better in fifteen, the Queensland children in only one, and in three the facility values were the same.

In interpreting these findings it is necessary to remember that significantly lower scores were obtained by the Australian children in the reading comprehension test. Thus it is possible that in answering the more complex verbal problems, the Queensland children were handicapped to some extent by difficulties in reading comprehension. Alternatively, it is possible that the children from this country, because they had been prepared for tests of this kind, were more practised in answering verbal problems of this type. Whatever the explanation may be, the results obtained from the two arithmetic tests indicate that, despite the near equivalence of the mean scores, the Australian children are better able to perform many of the basic operations of arithmetic than children of the same age in England and Wales.

---

[1] The numbers in the items quoted here have been changed slightly.

It is now possible to make a brief comment on the reasons for the markedly different distributions of test scores obtained by the Queensland children. It is clear that the relatively small proportion of children from Queensland scoring outside the approximate range of ± one standard deviation from the mean on the arithmetic tests (Table 2) is due in part to the content of these tests. They were made for children in this country and therefore included a number of items that were inappropriate for the Australian children. Nevertheless, it may be argued, that had these inappropriate items (fractions and decimals in the mechanical paper, and complex verbal problems in Arithmetic Test 2) been omitted, although the proportion of Queensland children scoring above 114 would have increased, the proportion scoring below 85 might well have decreased, giving further evidence of the superior performance in arithmetic of the Australian children.

Had the contents of these two arithmetic tests been equally appropriate for children from both countries, it might have been argued that the relatively greater dispersion of scores obtained in England and Wales was due to the system of streaming that is fairly widely employed—because of the demands of the 11 + examination "A" stream children are often given the best teachers and are encouraged to cover the syllabus at a rapid rate, whilst " C " stream children tend to be given less experienced teachers and have little incentive to progress. The detailed analysis of the arithmetic tests, however, suggests that this is not the complete explanation, and this conclusion is supported by the results on the reading test. Although the proportion of " good readers " (above 114) in this country is greater, suggesting better teaching for the abler pupils, the proportion of poor readers (below 85) is slightly less than in Queensland.

It appears, therefore, that the results of the reading and arithmetic tests taken together offer no conclusive evidence for the superiority of either system of organisation. The relatively larger number of " good readers " and " good arithmeticians " in this country may in part be due to the greater pressure brought to bear on "A" stream children of this age because of the 11 + examination. At the other end of the scale, however, there are inconsistencies between the subjects tested. Arithmetic appears to be better taught in Australia than in England and Wales in that there are fewer low scorers on the tests, but the same cannot be said for reading. The differences in performances on the attainment tests may well be due, therefore, to the methods and approach employed in teaching the different subjects and not to any overall differences in the organisation of classes.

There still remains one fact worthy of comment. Although it might be possible to explain the observed differences in the attainments of the children from the two countries in terms of differences in the effectiveness of the instruction they had received, such an explanation could scarcely be valid with regard to the results of the non-verbal test of ability. Nevertheless, there is evidence, from other sources, to show that performance in tests of this kind is affected by cultural influences, and it is reasonable, therefore, to suppose that the differences found in the results of this test, as well as those in the attainment tests, may be ascribed to environmental factors.

(b) Comparison between England and Wales and California[1]

Only one test is being considered here—the timed arithmetic Test 2—and since only a proportion of the items in this test could be regarded as being comparable it has not been possible to use standardised scores in the analysis. The results which follow, therefore, are given in terms of raw scores. It will be recalled that the number of strictly comparable items was 70, and the number of children contributing to the results was 3,191 from 91 schools in England and 3,179 from 70 schools in Central California. Table 8 gives the means and standard deviations of the scores from the two countries.

Table 8

MEANS AND STANDARD DEVIATIONS OF COMPARABLE ITEMS IN AN ARITHMETIC TEST GIVEN IN CENTRAL CALIFORNIA AND IN ENGLAND AND WALES

	California		England & Wales		Difference of Means
	Mean	S.D.	Mean	S.D.	
Section 1 Computation (31 items)	4·2	2·1	14·1	8.7	9·9
Section 2 Problems (39 items)	7.8	5·1	15·0	10·5	7·2
Total Test (70 items)	12·1	6.8	29·1	18.7	17·0

All the differences in Table 8 are highly significant.

As will be seen from Table 8 the differences between the two groups of children are very marked. A comparison of the distributions of total score emphasises the differences; the scores for the English and Welsh children spread over the entire range of the test from zero to a perfect score of 70, whereas there were no scores above 53 for the pupils from California. Thirteen pupils in California made scores above 38 as compared with 1,077 pupils in England who scored above this point.

From information obtained from California it was clear that this test, constructed as it was for children in this country, laid more emphasis in Section 1 on denominate numbers than is usual in American tests. An item analysis was therefore carried out which shows the percentage answering each of the 70 items correctly. The results are given in Table 9 and in order to complete the comparison the relevant figures for Queensland, Australia have been included with those from California and England and Wales. The item analyses were based on a random selection of 456 test papers from the English and Welsh sample and 544 from the Californian sample. The Queensland figures were obtained from the sub-sample of 150 children previously described.

---

[1] An account of this part of the investigation has been published in America. See " A Comparison of Achievement in Arithmetic in England and Central California," by G. T. Buswell, *The Arithmetic Teacher*, February 1958.

Table 9

PERCENTAGES OF CHILDREN FROM CALIFORNIA, QUEENSLAND AND ENGLAND AND WALES ANSWERING CORRECTLY EACH OF SEVENTY COMPARABLE ITEMS FROM ARITHMETIC TEST 2

a=California. *b*=Queensland. *c*=England and Wales.

SECTION 1 (MECHANICAL)							
Item No.*	Per Cent. Correct			Item No.*	Per Cent. Correct		
	a	*b*	*c*		*a*	*b*	*c*
	%	%	%		%	%	%
1	85	91	93	19	22	61	43
2	80	87	84	22	0	58	45
3	67	80	76	23	1	59	44
4	59	91	78	25	3	48	42
5	7	65	61	27	0	46	37
6	66	83	70	28	0	39	31
7	4	61	59	30	0	33	24
8	0	69	58	31	0	35	26
9	2	70	64	32	0	28	22
10	3	65	49	33	0	21	22
11	2	57	48	34	1	16	17
12	1	71	57	35	0	19	18
14	4	65	62	36	0	15	15
16	0	60	49	37	0	10	14
17	2	71	57	39	4	5	9
18	1	63	47				

SECTION 2 (PROBLEMS)							
Item No.*	Per Cent. Correct			Item No.*	Per Cent. Correct		
	a	*b*	*c*		*a*	*b*	*c*
	%	%	%		%	%	%
1	89	81	91	32	7	43	47
2	85	91	84	33	9	41	45
3	65	87	73	34	7	33	34
5	66	37	59	36	19	30	33
6	55	87	69	37	17	40	31
9	23	48	54	41	4	29	29
12	42	61	46	42	7	17	30
13	15	27	61	44	7	24	21
15	13	68	37	45	3	17	24
16	24	40	48	46	5	21	23
17	19	39	51	47	3	13	22
18	4	23	34	48	7	9	25
20	21	43	43	49	1	4	12
21	7	58	40	52	5	7	12
22	33	32	47	55	1	3	11
23	39	73	52	56	1	9	21
24	28	44	44	57	1	5	13
25	7	53	45	59	1	2	9
29	7	43	47	60	1	1	3
30	6	35	34				

*Item numbers correspond to the numbers on the printed test form.

Table 9 shows that the percentage of English and Welsh children answering the item correctly exceeded that for the American children on all but two items, (Nos. 2 and 5 of Section 2). There were five items in the test which involved hundredweight. Although this term is included in the tables of the Californian State textbook there was no indication that it had been taught, and furthermore it is seldom used in America. The average per cent. correct on these five items, (Nos. 22, 27, 35 and 36 in Section 1 and No. 21 in Section 2) is 31 per cent. in England and Wales and 1.4 per cent. in California. If these five items were omitted from the test it would, however, reduce the mean score on the total test for English and Welsh children by approximately 1.6 points but would make no significant difference in the comparison.

Table 10 below gives some examples of the items employed in the test with the percentages from the three groups of children answering them correctly.

Table 10

SAMPLE ITEMS FROM ARITHMETIC TEST 2 WITH PERCENTAGES OF CHILDREN FROM CALIFORNIA, QUEENSLAND AND ENGLAND AND WALES ANSWERING CORRECTLY

ITEM NO.	ITEM*	PERCENT. ANSWERING CORRECTLY		
		Cal.	Q'land	E. & W.
SECTION 1				
1	Add: 4 row t.u. with carrying figures	85%	91%	93%
3	Multiply: h.t.u. by 6	67%	80%	76%
4	Divide: h.t.u. by 7	59%	91%	78%
5	Add: 3 row ft. in.	7%	65%	61%
10	Subtract: gal. pt.	3%	65%	49%
12	Divide: lb. oz. by 9	1%	71%	57%
30	Multiply: hr. min. by 6	0%	33%	24%
SECTION 2				
2	John has 7 sweets and I have 3 times as many. How many have I?	85%	91%	84%
5	Underline the smallest fraction: $\frac{1}{3}$, $\frac{1}{10}$, $\frac{1}{8}$, $\frac{1}{6}$, $\frac{1}{4}$	66%	37%	59%
12	Father was 32 years old in 1953. When was he born?	42%	61%	46%
13	What is half of $9\frac{1}{2}$?	15%	27%	61%
15	Write down the number with eighteen hundreds and five units	13%	68%	37%
25	How many pints of water can be drawn from a 22-gallon barrel?	7%	53%	45%

*The numbers in the items quoted here have been changed slightly

Tables 9 and 10 illustrate further the slightly superior performance in arithmetic of the sample of children from Queensland, Australia, over their counterparts in this country; they also illustrate the markedly inferior performance of the sample of children from California, U.S.A., and it is interesting to examine some of the factors that may have contributed to this latter result.

It is, of course, necessary to appreciate that there are a number of fundamental differences in the educational systems of the two countries any or all of which may have had some part to play in the differences observed in this study.

(*i*) *Curriculum Differences*

In this country the State plays little part in determining the curriculum to be followed by its schools, but California, like other American States, prints its own textbooks and largely dictates what is taught grade by grade. The age at which the arithmetic appropriate to each of the seventy items of the test was taught can therefore be determined. One item appeared in the Californian textbook for Grade III (i.e. 8-9 years); nine in the book for Grade IV (9-10 years); thirty in the book for Grade V (10-11 years); and thirty in the book for Grade VI (11-12 years). Since some of the children in the age-range considered were still in Grade V they had not yet studied some of the items, and hence could not be expected to get them right. It is not possible to determine in a similar way the exact age at which the items were introduced to British children, since most of the textbooks available in this country do not state the age at which arithmetical processes should be taught. However, an analysis of the Beacon Arithmetic Books, which do suggest appropriate ages, reveals that only one item appears at the same age in both countries, thirty-seven items appearing one year earlier in the English textbook and thirty-two items two years earlier.

One other important difference in the curricula of the two countries should be noted. Here the basic arithmetic syllabus is completed at the age of eleven and children move on to secondary schools. American children continue their elementary, (primary) education through grades 7 and 8 and do not begin secondary school until the age of 14. " But," according to Professor Buswell, " while American children are finishing arithmetic in grades seven and eight, English children of corresponding age who enter their grammar schools are covering two years of high school mathematics. The issue is whether England puts too much pressure on pupils or whether California delays its program too much."

(*ii*) *Grade Placement v. Age Grouping and Streaming*

Most of the children attending schools in England and Wales are streamed for ability within each year group, and the brighter children are encouraged to proceed at a faster pace than those in the " B " and " C " streams. In addition, as has already been noted, it is not an uncommon practice for the more efficient teachers to be assigned to "A" streams and the relatively inexperienced ones to the "C" streams. In the American schools all children within each grade move at very much the same pace although those considered not to have covered the ground of one grade may repeat it for a further year instead of being promoted. It might appear, therefore, that the British system would necessarily result in a better average performance, but it must be remembered that the system of grade placement is also employed in Australia.

(*iii*) *Age of beginning school*

English and Welsh primary schools admit pupils at age 5 whereas the customary age for the first school grade in America is 6. However, over 97 per cent. of the pupils in the Californian sample had attended a kindergarten admitting five-year-olds. The differences in performance cannot, therefore, be attributed to length of schooling, although the way in which the early school years are used may be a contributory factor. Little, if any, formal teaching is carried out in American kindergartens whereas in this country some formal instruction in reading and number is given to many five-year olds.

(*iv*) *Teacher training*

The Californian certificates of teaching require a four year college course of which approximately four-fifths is liberal education and one-fifth professional education. Thus, in terms of academic training the Californian teachers have had more years of preparation than most primary school teachers in this country. Elementary (primary) school teachers in the United States, however, suffer peculiar limitations in the case of arithmetic, since the subject is usually not taught in high school nor is it part of the mathematics syllabus in college. Hence, except for short review courses sometimes provided by teachers' colleges, elementary teachers have comparatively limited advanced training in teaching the subject. As a result, the specialised study of arithmetic for most American teachers does not go far beyond what they received in the last year of the elementary school—which is not very much more advanced than that taught to eleven-year-old primary school children in this country. It seems fairly clear, therefore, that teachers in this country probably have a considerably better background in arithmetic than do teachers in California.

(*v*) *Different Social Conditions*

While it may be clear that the factors considered so far contribute in varying degrees to the observed differences it is necessary in reviewing the evidence to consider two other facts which might have an adverse effect on arithmetic attainment in California. In the first place there is a fairly high degree of mobility of the population in California compared with England and Wales, due in part to migratory agricultural workers, in part to the prevailing tendency to move frequently from one location to another and in part to the continuous and large influx of population from other parts of the country. In two of the larger schools included in the sample the change in school enrolment during the year amounted to more than 100 per cent. A second possible adverse factor is that of race and nationality. In two of the larger schools included, more than 98 per cent. of the pupils were negroes. The mean scores on the seventy items for these schools were 6.2 and 6.6 compared to the mean of 12.1 for all schools in the Californian sample. Also, judging as nearly as possible from the names of the children, there were 211 Mexicans in the sample who made a mean score of 9.3.

VI. Summary and Conclusions

A national survey of basic attainments was carried out in England and Wales in May, 1955. One of the age groups concerned consisted of children aged between 10 years 9 months and 11 years 8 months at the time of testing. The tests used were specially prepared tests of arithmetic, reading comprehension and non-verbal ability.

These tests were applied in November, 1955, to a comparable sample of children in Queensland, Australia.

A modified form of one of the tests of arithmetic was also administered in January, 1957, to a comparable sample of children in California, U.S.A.

The results of the surveys have been compared and the following is a summary of the major findings.

1. The performance of the English and Welsh children was found to be superior in both reading and non-verbal ability to that of the Queensland children. (The latter are regarded as average in these respects in relation to other Australian children, but as somewhat superior in arithmetic.)

2. No significant difference was found between the mean scores of the two groups in the tests of arithmetic.

3. A closer inspection of these results, however, revealed a number of interesting differences.
   (*a*) Few of the Queensland children had been introduced to decimals or fractions and were, therefore, at a disadvantage in attempting some of the items in the tests.
   (*b*) The Queensland children's performance was also poorer than that of the children in this country in the more complex problems. This is probably related to the difference in levels of reading comprehension.
   (*c*) The Queensland children were also shown to be at a disadvantage in the test which involved strict time limits and required a fast rate of working. This difference may perhaps be ascribed to the fact that many children in this country, as part of their preparation for the eleven plus examination, are practised with speeded tests.
   (*d*) Taking these factors into account the analysis showed that the Queensland children were markedly superior in mechanical arithmetic involving the application of the four rules, conversion units, etc.

4. The standard deviation or " spread " of the scores obtained by children in this country was larger in all the tests. This may in part reflect the effects of " streaming." It is sometimes alleged that one of the results of streaming is that the brighter children make more rapid progress than they would otherwise achieve, and also that the duller children, when assigned to a " C " stream, become more backward, partly as a result of poor morale, and partly because there is a tendency in some schools for the more experienced teachers to be put in charge of the abler streams. It should be noted, however, that as far as reading is concerned, there were fewer " poor " readers found in the English and Welsh sample than in the Queensland group.

5. The comparisons with the children in California were based on one of the arithmetic tests only. (This was modified somewhat because some of the items—especially those involving money sums—were not applicable in American schools.) The children in Central California are regarded as about average in attainment in relation to children in the United States as a whole.

   The performance of the Californian children was found to be markedly inferior to that of children in England and Wales in both the mechanical and the problem arithmetic items.

Although children in California do not enter the elementary schools until 6+, the majority of the sample had in fact attended kindergarten schools. The main reasons for the different levels in performance are probably the fact that formal teaching tends to be introduced at an earlier age in this country, and the fact that there is a difference in the standards in the two systems. In the United States more limited objectives are formulated for children of primary school age, less emphasis being placed on rapid progress in mechanical arithmetic than is customary in this country.

It is interesting to note that Queensland and California both have the grade system of organisation as distinct from the form of organisation prevalent in this country whereby children are promoted by age and, in the majority of primary schools, streamed or grouped according to ability. Nevertheless, the results show that Queensland children tend to be superior, Californian children inferior to English and Welsh children in arithmetical attainments. This would suggest that the form of organisation is not the most important factor in determining levels of attainment. The most important factor would seem to be the definition of standards and objectives and the amount of time and effort that it is considered desirable to devote to any particular aspect of the curriculum. No doubt, there are also differences, as the comparison between California and this country suggested, between the kinds of training and experience that teachers enjoy and their consequent effectiveness, which may have a bearing on the standards achieved by their pupils.

Abstract

FACTORS RELATED TO READING DISABILITIES IN THE FIRST GRADE OF THE ELEMENTARY SCHOOL

Doctor Thesis, University of Stockholm, 1958. *Acta Universitatis Stockholmiensis, Stockholm Studies in Educational Psychology* 2. *Stockholm: Almquist and Wiksell.*

by Eve Malmquist

THIS study deals in a most thorough and comprehensive way with the early stages of learning to read and the causes of reading disability in children of 7-8 in Swedish schools, that is, at the end of their first school year. It grew out of the author's experiences as reading teacher and psychologist in the municipal primary schools of Malmo and from an analysis of the results obtained in over 350 (cited in Bibliography) Danish, English, German and Swedish studies.

Purposes

The main purposes of the investigation were limited to the following: (i) to examine critically the concepts of reading disability so far used, and to suggest an operational definition capable of exact description; (ii) to produce the necessary measuring instruments for selecting children in need of remedial treatment; (iii) to clarify the relationship of certain factors with reading disabilities at the end of the first grade; (iv) to analyse the types of reading and spelling errors made at this stage; and (v) to see how far such disabilities tend to persist throughout the primary school course.

Method

To answer these questions the author constructed and standardised a battery of five tests covering different aspects of oral and silent reading, two spelling tests and five tests of visual perception. These instruments were of satisfactory reliability and validity and compare favourably with similar tests in use elsewhere. The battery of visual perception tests which includes figures, letters, digits, geometric, nonsense and various other configurations is particularly thorough and ingenious. Malmquist also developed rating scales for those variables which are not readily susceptible to quantitative measurement, e.g. personality traits, and collected other relevant data from interviews and various official sources.

Definition of 'Good' and 'Poor' Readers

Examination of the many concepts of reading disability put forward by medical and educational writers revealed that for the most part these are based upon unverified causal assumptions or are too vague. The author defined his concepts in meaningful statistical terms by reference to his instruments of measurement. Briefly, 'poor' readers are those children whose reading index (combined scores on his five nationally normed reading tests) falls at or below one sigma from the mean. 'Good' readers correspondingly have scores at or above one sigma from the mean, and ' Medium ' readers are those who lie within the limits of plus or minus one sigma from the mean. A sub-group of children with ' special ' reading disability is also distinguished, consisting of pupils whose reading index is one sigma away from their intelligence quotient.

Sampling

For the main investigation random samples (homogeneous in respect of age and residence) were taken from all first-grade pupils in two school districts: Kristianstad and Linkoping. These numbered 399 children aged seven to eight who were divided into groups of ' poor ', ' medium ' and ' good ' readers according to the criteria, and assessed with regard to the more significant physical, intellectual, social and developmental factors which have been adduced as associated with reading ability.

Each factor was studied separately in its relation to reading disabilities but, as a supplement to the more traditional methods of investigation, an attempt was made to study also the interactions of several variables at a time by using the method of analysis of variance of higher order (multifactorial design).

Results

(i) *Ability*

Statistically, significant differences existed between poor and good readers with regard to intelligence (as measured by the Terman-Merrill individual test scale).

(ii) *Personality traits*

All but one (the ' need for social contacts ') of the personality traits considered significantly differentiated good from poor readers. Thus, besides being generally less intelligent, poor readers tend to have less self-confidence, less persistence and concentration, less ability to make contacts, and to be less stable and more inclined to submit to the leadership of others than are children of good reading ability.

(iii) *Socio-economic variables*

Of the socio-economic and socio-cultural factors studied, parental social status, joint taxable income, and education above the elementary school level, as well as the number of rooms and books in the home were all strongly associated with good reading in children. ' Broken ' homes, family size, and a child's ordinal position in it, were not shown to be significantly related to reading ability.

(iv) *Pre-school history*

Consideration of the children's pre-school development revealed that premature birth and an exceptionally low birth weight were more common among poor readers, but that birth injuries, speech defects, severe diseases and non-attendance at a nursery school were not significantly associated with either group.

(v) *Physical and physiological variables*

Left-handedness and defects of hearing and vision were not found to be related to poor reading ability. However, mean differences between poor and good readers in the scores on the visual perception tests were significant in every case. It therefore seems likely that further work on this test battery would be useful diagnostically.

(vi) *Teacher variables*

A significant relationship between the teacher's age, number of years' service and the pupils' achievements in reading was revealed, but the teacher's knowledge and efficiency as indicated by his training college diploma and final teaching mark bore little relation to them.

(vii) *Absence*

The hypothesis that poor readers are absent during the first school year to a greater extent than medium and good readers was not verified in this investigation.

(viii) *Sex differences*

An analysis of sex differences revealed that the girls obtained somewhat better results than the boys in spite of their lower average ability as measured by the Terman-Merrill. No significant differences were obtained in most of the other variables except that among girls there were fewer premature births and teachers rated them higher in respect of the personality characteristics which were shown to differentiate strongly between poor and good readers, i.e. persistence, concentration and emotional stability.

(ix) *Special reading disability*

Children with ' special ' reading disabilities in the sense that their reading ability was very low (below one sigma on the distribution curve for the reading tests employed) but whose intelligence was normal or supernormal (I.Q. 100 or above on the Terman-Merrill scale) occurred in 4.8 per cent. of the whole combined sample, and when the line of delimitation for normal intelligence was drawn at an I.Q. of 90 this percentage was increased to 8.5. Pupils whose reading ability deviated significantly (more than one sigma) in a negative direction from general intelligence occurred in 20.8 per cent. of the whole combined sample.

Using the case-analysis approach, a separate study of children with reading difficulties and an I.Q. of at least 90 (34 cases) was also made. This showed that reading disabilities at first-grade level are never isolated effects, for all such pupils deviated negatively, in a very marked manner, from the mean for the total population investigated with regard to the major factors studied.

(x) *Persistence of difficulties*

In order to get an idea of the degree of consistency of reading disabilities and also to ascertain the predictive value of the tests used, a follow-up study of all but one of the original samples of children was made when they were in grade four. Correlations between the results of reading tests given in the first grade and marks obtained by the children in the fourth grade were all above .6. Furthermore, with only one exception, all pupils designated as poor readers at the beginning of the investigation remained at a level of achievement below the average of that of their school fellows of the same age in the populations investigated.

(xi) " Word Blindness "

One other finding is of importance since it appears to dispose of the misleading notion of ' word blindness '. Except for the omission of vowels in spelling, all the various alleged symptoms of this ' disease ' entity were found in readers at all levels although to a greater extent among subnormal readers. The author therefore concludes that differences between poor and good readers are of a quantitative rather than a qualitative nature.

Conclusion

This study combines scientific objectivity with such a wide scope that its contributions to knowledge are many and varied. Its wide representative sampling from a normal population is a departure from the usual method adopted by research workers of selecting children from reading clinics, remedial reading classes, and child guidance clinics. In this way a better opportunity is afforded for studying representative groups of poor readers and children with ' special ' reading disabilities. The application of operational methods for the differentiation of various groups of readers in conjunction with these sampling principles also makes it possible to obtain an unequivocal and clear definition of the incidence of reading disabilities in any population.

Besides its methodological value this study makes a direct contribution to an important area of applied educational psychology and child development. It strongly supports the hypothesis of multiple causation and the notion of reading disability as a phenomenon within the normal range of variation.

J.M.M.
W.D.W.

A SELECTED AND ANNOTATED BIBLIOGRAPHY

J. B. Biggs

Explanatory Notes

The inclusion of a particular book or article in the bibliography which follows indicates that, in the compiler's view, it is worthy of consideration by teachers. The fact that a given book does not appear does not necessarily imply that it has been rejected, but simply that it may have escaped the notice of the compiler.

The publications listed below fall into three categories. The first is mainly concerned with the teaching of arithmetic in primary schools, although secondary education is covered to a limited extent; the second deals with books which put forward somewhat less orthodox approaches to the teaching of number, and with general books from overseas, notably the United States; the third covers reports of research and books on the psychology of arithmetic.

A caution should be added with regard to those books in Section II which advocate new methods of teaching. While each method has been tried out in at least one school, there has not been, as yet, any controlled research definitely establishing its worth, or otherwise. This is not to say that new methods are of no immediate value, but simply that one must exercise due care in accepting their authors' claims. While awaiting objective validation, it is, nevertheless, refreshing—and important—to be at least aware of different approaches to the study and teaching of arithmetic.

The asterisks (*) beside each title give an indication of the level of difficulty in terms of the background knowledge of statistical or psychological theory, etc., presupposed by the author.

*—Easy. **—Moderate. ***—Difficult.

IMPROVEMENTS ON CURRENT METHODS

***Arithmetic in Primary Schools.** ASSOCIATION OF TEACHERS IN COLLEGES AND DEPARTMENTS OF EDUCATION. Longmans, Green & Co., London, 1954. Sixth impression, 76 pp.

This is a report on the teaching of mathematics in primary schools in England. The purpose of the book is to supplement American reports. The writers try to show how the content of the subject, as distinct from method, may be enriched; how the content and suggested standards may be evaluated; and how the curriculum may be evaluated in the light of the A.T.C.D.E. Committee's observations, and available research. One interesting view the committee puts forward is that ' spatial knowledge ', or informal geometry, should be taught from the earliest stages, replacing a proportion of the purely arithmetical curriculum.

***Romance in Arithmetic.** M. E. BOWMAN. U.L.P., 1950. 160 pp.

A book containing illustrations and amusing examples designed to relieve the drudgery involved in mastering the basic number facts and processes.

***Traditional Number Rhymes and Games.** F. DOREEN GULLEN. Scottish Council for Research in Education, No. XXXII. U.L.P., 1950. 143 pp.

This collection of rhymes and jingles was assembled at the request of the Scottish Council's panel on early Number teaching. The aim was two-fold: to provide infant teachers with informal number material, and to try to recover the dying heritage of Scottish number rhyme. While many of the rhymes are in a rich Scots dialect, there are others which should be very suitable for the infant teacher in English schools.

****The Principles of Arithmetic.** A. KEITH and J. ROBERTSON. Blackie, London, 1949. 278 pp.

This book originates from lectures given by the authors to Training College students, and is directed in the main to Junior and Secondary Modern teachers. It is the authors' belief that " the teacher should have a knowledge of arithmetic far in excess of the standard to which he will be required to teach the subject " and the tone of the book is in keeping with this end. The reference to past research and the close analyses of methods of teaching the four rules are extremely useful.

***The Teaching of Arithmetic in Primary Schools.** K. LOVELL and C. H. J. SMITH. U.L.P., 1953. 128 pp.

This is a practical book addressed largely to Training College students and newly qualified teachers. It covers most aspects of the Infant and Junior curriculum. As in the first book mentioned, these authors recommend that more attention should be paid to ' spatial knowledge '. Other recommendations include such practical advice as how to get the best out of chanting tables, the equal-additions method in subtraction, and, in general, as much pre-, extra and intra-school number experience as possible.

***The Teaching of Mathematics in Post-Primary Schools.** J. H. MURDOCH. Studies in Education, No. 12, N.Z.C.E.R. O.U.P., 1950. 183 pp.

Mr. Murdoch writes primarily for New Zealand Secondary teachers but nevertheless, a great deal of what he has to say about the general problems of secondary mathematics—particularly as it concerns Secondary Modern teachers—is erudite, practically sound and highly readable. Newcomers to the teaching profession will find Mr. Murdoch's maxims most valuable.

***Fundamental Number Teaching.** R. K. and M. I. R. POLKINGHORNE. Harrap, 1955. 196 pp.

This book was written in response to the requests of 2,000 teachers for advice from *The Teachers' World*, and has thus a directly practical aim. The authors start from the very first number experiences and work right through the Primary School curriculum. With Schonell, the authors recommend that basic addition facts should be learned systematically, in the same way as, sooner or later, the child must master the multiplication tables. Chapters especially worthy of note are those on place value and zero, and on methods of subtraction.

***Diagnosis and Remedial Teaching in Arithmetic.** F. J. and F. E. SCHONELL. Oliver and Boyd, London, 1957. 195 pp.

This book supersedes the authors' previous book, *Diagnosis of Individual Differences in Arithmetic*, and covers more ground in the light of the most recent research. In the first part of the book the authors consider the factors making for good achievement in arithmetic, what research has to say on arithmetic teaching, and the various causes of backwardness. Then follows a discussion on diagnosis of difficulties in arithmetic, and the way in which the Schonell Diagnostic Arithmetic Tests can assist diagnosis, and finally, a programme of remedial teaching on the basis of the Schonell test results. Thus, the book is of great importance to teachers in two ways: it outlines the most common difficulties encountered by pupils in learning the fundamentals of arithmetic; and it suggests methods for diagnosing and remedying the difficulties of individual pupils.

***Number.** THYRA SMITH, BLACKWELL. Oxford, 1954. 188 pp.

***A Background of Primary School Mathematics.** L. D. ADAMS. O.U.P., 1953. 189 pp.

These two books are written by retired H.M.I.'s both of whom have had considerable experience as teachers and inspectors. On the basis of this experience, they offer eminently readable impressions and suggestions of interest to Primary teachers and parents. The sections devoted to suggestions for widening the scope of informal number work are of particular value.

***The Teaching of Arabic Numerals.** G. G. N. WRIGHT. S.C.R.E., No. XXXVI. U.L.P., 1954. 45 pp.

This is an interesting history and analysis of the construction and form of Arabic numerals. The author analyses ambiguities and mistakes commonly made by adults and infants in the formation of numbers, and suggests a simplified and less ambiguous technique for teaching the writing of number. It should be of interest especially to the Infant Teacher, and also to those who teach handwriting and art.

*****The Mathematical Association** from time to time issue reports (published by G. Bell & Sons) specifically addressed to teachers and almost entirely concerned with the teaching of mathematics. These reports are very detailed and thorough, and should afford considerable assistance to the teacher in the planning and presentation of his subject. The journal of the Association is *The Mathematical Gazette*, issued four times a year, and is available to non-members. The post-war publications are as follows:

INTERIM REPORT ON THE TEACHING OF MATHEMATICS IN SECONDARY MODERN SCHOOLS; No. 29, 1949. (Out of print.)

REPORT ON THE TEACHING OF MATHEMATICS IN SECONDARY TECHNICAL SCHOOLS; No. 30, 1949. (Out of print.)

REPORT ON THE TRAINING OF TEACHERS OF MATHEMATICS IN TECHNICAL SCHOOLS AND COLLEGES; No. 31, 1949.
REPORT ON THE TEACHING OF TRIGONOMETRY IN SCHOOLS; No. 32, 1950.
REPORT ON THE TEACHING OF CALCULUS IN SCHOOLS; No. 33, 1951.
REPORT ON THE TEACHING OF HIGHER GEOMETRY IN SCHOOLS; No. 34, 1953.
THE USE OF VISUAL METHODS IN TEACHING MATHEMATICS; No. 35, 1954.
A LIST OF BOOKS SUITABLE FOR SCHOOL LIBRARIES; No. 36, 1954.
REPORT ON THE TEACHING OF MATHEMATICS IN TECHNICAL COLLEGES; No. 37, 1955.
REPORT ON THE TEACHING OF MATHEMATICS IN PRIMARY SCHOOLS; No. 38, 1955.

II. THE TEACHING OF ARITHMETIC—*Less Orthodox Approaches.*

***Counting and Arithmetic in the Infants School.** W. G. BASS and O. S. DOWTY. Harrap, London, 1956.

This is an account of a method of teaching arithmetic, which has been tried out in an infants' school. The book deals chiefly with infant work, describing in detail the method and apparatus used. The method is based on the proposition that " Arithmetic becomes largely an exercise in translation—from formal activities and words into symbols and signs . . ." The means by which this ' translation ' is carried out is based entirely on counting. Numbers are shown to be counting symbols and the different ways of counting are related to the four rules—each process being a logical development of the preceding one. The novelty of the approach to the subject should be of great interest to all infant teachers.

***Arithmetic in Action.** E. BRIDEOAKE and I. D. GROVES. U.L.P., sixth impression, 1954. 128 pp.

This is the report of an experiment using the ' realistic ' approach. Life situations are replicated on a smaller scale for use in the Infant Department, and constitute the entire formal number experience of the infants concerned. Full details of the method and the easily constructed apparatus are given. The main claim of the authors is that with this approach the children develop a far more favourable attitude to arithmetic and without any lowering of the academic standards eventually attained.

****How to make Arithmetic Meaningful.** L. J. BRUECKNER and F. E. GROSSNICKLE. John C. Winston Co., Philadelphia, 1947. 513 pp.

" The major objectives of the modern arithmetic program are: (1) to develop in the learner the ability to perform the various number operations skilfully and with understanding; and (2) to provide a rich variety of experiences which will assure the ability of the pupil to apply quantitative procedures effectively in social situations in life outside the school " (p. 1). Pages 2-513 of this book then proceed to develop the second of these themes. A great deal of research is reported in the book, but it is almost entirely related to U.S. conditions and hence needs intelligent adaptation to this country. Chapters of interest are those on ' The Modern Arithmetic Program ' and ' The Changing Arithmetic Curriculum ', Chapters V to IX, covering the first steps in arithmetic to decimal fractions, might be immediately useful. The last four chapters cover such topics as testing, remedial work, problem solving and apparatus.

****The Elementary School Subjects** (Part III: Arithmetic). LUELLA COLE. Rinehart & Co., Inc., New York, 1946, pp. 321-448.

This publication, like that of Brueckner and Grossnickle above, demonstrates the difference between British and American attitudes, particularly to the content of the syllabus. There are topics treated here, however, which are not usually dealt with in British texts, and these are of definite relevance—for example, the chapters on the analysis of difficulties with vulgar fractions and the solution of verbal problems.

****Numbers in Colour.** G. Cuisenaire and C. Gattegno. Heinemann, London, 1954. 39 pp.

This book explains the method (originating in Belgium and in growing use elsewhere of teaching arithmetic to primary school children by means of Cuisenaire's carefully constructed system of coloured rods. Taking 1 cm. as a convenient unit length of rod, the remaining numbers are treated in ' factor families ', each family (e.g. numbers having a factor of 3) having an allotted range of colour. The idea is that by this means, the child will associate an extrinsic property (colour) with an intrinsic and necessary property of number, and thus, by experiment and active manipulation, he will acquire a ready understanding of number concepts and the four rules of number. This apparatus has been compared with the American Stern material, the main differences being that the latter makes no use of colour relations, and is more expensive and complicated.

****Teaching the New Arithmetic.** G. M. Wilson, M. B. Stone and C. O. Dalrymple. McGraw Hill Book Co., New York, 1939.

This book reflects clearly the essentially pragmatic approach which characterises much of the education in the U.S. More specifically, these authors adopt the " more rational view of arithmetic as a functional service tool. . . . Common adult usage should be the limit of any arithmetic undertaken for drill mastery. Work beyond this . . . if recreational . . . can be left to choice . . . if vocational . . . can be left to learning on the job by the few who need it . . . or if encyclopaedic, it can be left in the encyclopaedia." The authors, on the basis of much sound research, outline those topics in mathematics which are likely to be of practical use in adult life; those which are not are to be rigorously excluded from the curriculum. While many in this country might not be prepared to go as far as these authors, particularly in regard to the implications for technical education, this is an interesting and refreshing book, the chapters on problem work especially so.

III. GENERAL RESEARCH AND THE PSYCHOLOGY OF ARITHMETIC.

***Problems in Mathematical Education.** H. Dyer, R. Kalin and F. Lord. Educational Testing Service, Princeton, 1956. 50 pp.

Following expressions of public dissatisfaction with the teaching of mathematics, the Carnegie Corporation asked the Educational Testing Service to describe the kinds of research activities likely to lead to better courses and teaching in mathematics. The result is this booklet, summarising past work in a bibliography consisting of 250 references, with general remarks as to the present implications and the suggested trend of future research. There are ruthless criticisms of training requirements, current attitudes to the subject and ' teacher readiness ', some of which might well apply to conditions in this country also.

****The Teaching of Arithmetic.** National Society for the Study of Education, 50th Year Book, Part II. University of Chicago Press, 1951. 308 pp.

This yearbook is a summary of the most recent and important research projects on the teaching of arithmetic. Naturally, the context is that of the American educational system, but much of what is said is relevant to English schools. The authors stress ' inductive ' and ' relational ' methods of teaching—an emphasis based more on theoretical grounds than on strictly experimental research. Chapters of more direct practical interest are those by Swenson Spitzer and Horn. There are other valuable chapters bearing on wider aspects of number teaching—for example, the mathematical background needed by teachers of arithmetic and suggestions for research.

*****Studies in Arithmetic.** Volume I, S.C.R.E., No. XIII. U.L.P., 1939. 144 pp.

This is the first progress report of the Scottish Council Sub-Committee on arithmetic. While it is true that much of the research is restricted in applicability and partly superseded by later investigations, nevertheless there is a good deal of work that is of continuing value and importance, particularly to the primary school teacher. Among the topics covered are: the time allotted to arithmetic, the relative difficulty of basic number facts (see also Publication Nos. XXVIII and XXXVII) and adding upward versus adding downward.

*****Studies in Arithmetic.** Volume II, S.C.R.E., No. XVII. U.L.P., 1941. 218 pp.

This is the second collection of number researches published by the S.C.R.E. The samples were restricted to Scots children, and the research was undertaken over 15 years ago, but the following chapters especially are recommended as being important for teachers: the relative merits of methods of teaching subtraction (all the evidence points to the superiority of the equal-additions method); zero and its attendant difficulties; analysis or errors in fractions; and finally, an intricate report on placement of topics in arithmetic, in which the findings of the Washburne Committee of Seven are criticised.

*****The Child's Conception of Number.** JEAN PIAGET, ROUTLEDGE and KEGAN PAUL. London, 1952. 248 pp.

This book is based on a series of experiments or test situations planned to investigate the genesis of number concepts from early infancy to adolescence, when the child should be ' arithmetically mature '. Piaget's main hypothesis, which he holds to be proven by his evidence, is that the construction of number concepts is concomitant with the development of logical thought, and that therefore logical and arithmetic operations constitute a single, psychologically natural, system. He sees the notion of ' constancy ' or ' conservation ' as being the primary basis of the psychology of arithmetic, as it is in the case of the physical sciences.

In his investigation of cardinal and ordinal quantities, he discovers three stages in all the operations considered. The first stage is based solely on perceptual judgments of ' greater ', ' equal ', ' smaller '; so that the four-year-old child does not readily perceive that the volume of water, for instance, remains constant when poured from a smaller to a larger container. The child then passes through an intermediate or transitional stage in which the visual percept may or may not be more powerful than the principle of constancy. Finally, around the age of seven years, the average child suddenly grasps the ' primacy of conservation '. The implications of this for the primary school teacher are evident. It is, for example, clear that until the child has reached this third stage he has little hope of comprehending the essential meaning of ordinary addition and subtraction, or even of measurement itself.

Criticism has been levelled at the possible lack of generality of Piaget's findings, at the ' loading ' of the outcome of his tests by the form of his questions. No-one, however, has seriously questioned his notion of sequence in development. The book represents an important contribution to psychological theory and may serve as a guide for the mathematics teacher who is prepared to translate its findings into practical terms.

For a useful summary of later developments in Piaget's work, in which number thinking is dealt with alongside other cognitive and affective aspects, the following is recommended:

*****Berlyne, D. E.,** " Recent development in Piaget's work." *British Journal of Educational Psychology*, Volume XXVII, February, 1957, pp. 1-12.

****The Early Development of Number Concepts.** S.C.R.E., No. XX. U.L.P., 1942. 63 pp.

This is the report of a survey conducted by the Scottish Council amongst Scottish Infant schools on the extent to which children have acquired number knowledge—ability to count, do simple sums, etc.—before they enter Infant school. There are also smaller sections on frequency of use, in Infant schools, of number games, pictures and finger counting.

*****Addition and Subtraction Facts and Processes.** S.C.R.E., No. XXVIII. U.L.P. 66 pp.

This report is a sequel to the preceding one, and concerns the child's knowledge of number on his transfer to the Junior school. 1,315 children, from 47 Infant schools, were tested. Although specifically related to the Scottish system, the results are certainly suggestive and important for English teachers. It is seen from the tables published that the relative difficulty of items tend to remain constant throughout the school year; and that, generally, as the size of digits increases from 0 to 9, so does the difficulty of the item. Numbers containing zero, however, are difficult at any level and tend to show less consistent gradation. By obtaining the same sorts of data about processes, it becomes possible to construct tests of pre-arranged difficulty and with a predictable distribution of marks.

*****Patterns of Error in the Addition Number Facts.** James M. Thyne. S.C.R.E., No. XXXVII. U.L.P., 1954. 275 pp.

Using the same data as the report on *Addition and Subtraction Facts and Processes* (S.C.R.E., No. XXVIII), Thyne takes the research an important step further by enquiring into the *nature* of error—the psychological processes involved—as distinct from mere frequency of kinds of error. The author makes the point that when the teacher understands " answer-producing " mental activities, he has much greater hope of remedying mental processes productive of error. The author proposes a tentative psychological theory of meaning, closely analyses the errors obtained in the light of this, and then suggests practical implications which follow for the teacher.

TEST ANNOUNCEMENTS

Two distinct series of tests are constructed at the Foundation. They are:

" Closed " Series

Standardised tests for allocation at 11+ and for transfer at 13+ are produced each year. A catalogue giving information about these tests will be sent to local education authorities on request.

" Open " Series

Tests in this series have been prepared specially for the use of the teacher in the classroom. Details and prices are given in the catalogue, " Educational Guidance in Schools: Standardised Tests for the Use of Teachers," obtainable from the Foundation.

In addition to the above tests constructed by the Foundation's staff, a large number of tests and test materials are distributed through the Foundation's *Test Agency* on behalf of the Psychological Corporation of America, the Australian Council for Educational Research and various authors in this country. A catalogue giving brief descriptions of the tests, their availability and prices, can be obtained from the Foundation.

The following tests have recently been made available through the Foundation's Test Agency:

Wing Standardised Tests of Musical Intelligence

These tests are designed to measure the children's mental capacity to cope with music. The results can be expressed in terms of gradings according to age. These are of considerable help in giving some guidance in the following directions:

1. (1) Picking out musically bright children who can profitably study a musical instrument, e.g. for the school orchestra.
2. (2) Selecting children who are capable of studying music up to G.C.E. levels, or students who might profitably follow an Advanced, Ordinary or basic course at a Training College.
3. (3) Dividing a group of school children, or of students, into 'sets' for the study of music.

The test consists of a number of standardised ear tests which are played on a piano and then recorded on tape. The tests include: Analysis of chords; Pitch discrimination; Memory for pitch; Harmony appreciation; Intensity appreciation; Rhythm appreciation; Phrasing appreciation. They are very easy to administer and are suited to a wide range of candidates of from 8 years to good adult status. Norms are based on many thousands of cases. The playing time of the test is about 45 to 50 minutes—say about an hour for testing students or two periods of about 40 minutes for testing children. Further details and prices are available from the Foundation.

A Battery of Diagnostic Performance Tests

With the publication of the book " Diagnostic Performance Tests " by Boris Semeonoff and E. L. Trist (Tavistock Publications, Limited, 1958) the following tests have now been made available through the Foundation's Test Agency:

1. The Semeonoff-Vigotsky Test.
2. The Trist-Hargreaves Test.
3. The Trist-Misselbrook Test.
4. The Carl Hollow Square.
5. The Passalong Test.

Further details and prices of these tests, and copies of the book, can be obtained from the Test Agency, 79, Wimpole Street, W.1.

© National Foundation for Educational Research in England and Wales.

Printed by King, Thorne & Stace Limited, 1-13 Jubilee Street, Brighton, 1,
and published by Newnes Educational Publishing Co., Ltd., Tower House, Southampton Street, Strand, London, W.C.2.,
The National Foundation for Educational Research in England and Wales, 79 Wimpole Street, London, W.1.

VOLUME 1. NO. 2

FEBRUARY, 1959

EDUCATIONAL RESEARCH

A Review for
Teachers and all concerned with
Progress in Education

Published for
THE NATIONAL FOUNDATION FOR EDUCATIONAL
RESEARCH IN ENGLAND AND WALES

by

NEWNES EDUCATIONAL PUBLISHING CO. LTD.
Tower House, Southampton Street, Strand, London, W.C.2

EDITORIAL

THIS second number of the first volume, apart from the important papers by Sir Cyril Burt and by Dr. C. M. Fleming, contains the first of a new series of articles devoted to the teaching of mathematics. Both this and the series concerned with reading are planned to continue intermittently over this and the next volume. They will not be confined in their attention to the primary school but some at least of the series will be devoted to such work as exists touching the teaching and learning of these subjects in all types of secondary school. Also in this number is an article on Urban-rural Differences which will be followed at intervals by others which are intended to discuss background factors affecting children's learning and progress in school or which should be considered when policy decisions on kinds of educational provision, curricula and teaching method, have to be made.

It is evidently not possible, even in an annual volume the length of a considerable book, to touch upon all subjects where worthwhile research has been done and on which those concerned with education need objective information. This journal, however, is intended to be a working instrument for our colleagues in schools, training colleges, universities and education offices. Editorial notions of priorities are not always an accurate reflection of real ones. We therefore shall be happy to have from our readers suggestions which will help us give the service desired; that is, we should like to hear of broad and special fields of interest on which research information may exist and which readers would like to know about. In some cases it may be possible, if organisers of courses and conferences can let us know well in advance for papers or bibliographies to be produced and published in the journal for those who are likely to attend. Each issue is, however, planned many months in advance and it would only be possible to do this if we get notice about six months before the issue in question is due to appear.

We should like to draw your special attention to two announcements which appear on the last pages of this issue. The first concerns the special binding case for which we have been asked, and which should enable those who wish to do so to keep their copies in good condition. The second concerns the series of occasional publications which is meant to supplement the other aspects of the N.F.E.R.'s Information Services by making available research material of limited appeal but high value.

Finally, may we thank those many people who have written and telephoned to wish the new venture well, and say how valuable they think the journal is and to make cosnstructive suggestions as to how it may be improved. We hope that you will continue to make constructive criticisms and suggestions.

GENERAL ABILITY & SPECIAL APTITUDES

by Cyril Burt

I. Definition of Intelligence

In the previous issue of this journal Professor Vernon has discussed the naïve and over-simplified notions of intelligence testing, still widely current among teachers and the general public; and his penetrating criticisms have done much to clear the ground. The editor has kindly invited me to summarize my own experience in this field; and my special task, I take it, is to try to rebuild, with what materials are now available, some kind of constructive theory on the vacant space Professor Vernon has thus swept clean.

Our present attempts to assess the abilities of school children by standardized tests have grown out of the pioneer work of Galton and Sully at University College some sixty years ago. It was they who introduced 'child guidance', by which they understood guidance for the normal and supernormal as well as for the subnormal; and it was the success of their methods which induced the L.C.C. to appoint an official psychologist to carry on their work in the London schools. Meanwhile, in Paris, Binet had already adopted Galton's plan of 'substituting objective measurement for subjective impression', and applied it to the 'diagnosis of children suffering from innate defects of the mind'. Binet's scale of tests and Galton's techniques for statistical evaluation are still in use today. Hence the first question to be settled is plainly this: what precisely were these investigators, and the educational psychologists who followed them, attempting to measure? After that, we shall have to ask: how far were their claims justified?

As a working model of the human mind Galton, Sully and Binet accepted the traditional scheme more or less as it was; but they introduced two crucial modifications to allow for the new Darwinian principles of heredity and individual variation. Reduced to its essential outlines, therefore, their picture might be compared to that of a two-storeyed building with a vertical partition separating the front and the back. The ground floor represents what is inborn or congenital, while the rooms above form a variable superstructure erected afterwards, and consisting of 'acquired attainments': in a later version Galton also mentions an 'underground basement', inhabited by unconscious impulses and memories, which are allowed admission to the 'presence-chamber' only when summoned by the servants' bell. At the rear of the building, not accessible to the ordinary observer, are housed the busy backroom operatives who keep things going—'the emotions and sentiments, the fears and the desires, that provide the real energy for the day's demands'. In front reside the 'intellectual organizers who supply the necessary guidance and direction': these form, as it were, a board of managers, and include a presiding chairman, whose function is the general supervision of all activities, and under him a whole hierarchy of subordinate officials—the 'special capacities or aptitudes'.

This threefold dichotomy thus culminates (and here lies the most original feature of Galton's scheme) in the conception of an all-important 'supervisory factor' which is defined by three essential attributes: it is an *intellectual* factor, as distinct from the emotional and moral; it is *innate* and not acquired; and it is a *general* capacity, and so is contrasted with the more specific 'faculties' recognized by the scholastic psychologists and the later phrenologists. For this super-faculty Galton's usual name is ' general ability ', but sometimes, when the argument requires, he speaks of ' natural ability '. In French the word ' habilité ' has acquired a somewhat different connotation; Binet, therefore, adopts the term ' l'intelligence '—a word coined by Cicero and revived by Spencer, whose *Principles of Psychology* had in those days a wide vogue especially in France. In this country ' intelligence ' still had a slightly alien ring. Teachers more often talked of the ' clever ' child; and academic psychologists, to escape the ambiguities that had gathered around the traditional names, preferred the nomenclature introduced by Stout. Instead of ' intellect ' (or intelligence), ' emotion ', and ' will ', they talked of the ' cognitive ', ' affective ', and ' conative ' aspects of mental life: those, for instance, are the terms employed by McDougall in his well-known description of the ' structure of the mind '.

Binet, in explaining how he came to develop his ' scale of tests ', leaves us in no doubt about what he intended to measure. First, he says, we must distinguish between ' intellectual development and character ': emotional or temperamental qualities, he thinks, can best be assessed by observation, not by experimental tests. For measuring acquired attainments he draws up a series of ' pedagogical scales '. And, when he comes to discuss innate capacities, he makes a clear distinction between ' general intelligence ' on the one hand, and ' partial aptitudes ', on the other: the latter, he says, ' will be the object of a later investigation '.

' What meaning then ', he asks, ' are we to give to this word so vague and comprehensive—*l'intelligence?* Nearly all the phenomena with which experimental psychology concerns itself are phenomena of intelligence: sensation, perception, memory, are intellectual manifestations as much as reasoning. . . . It seems to us there is one underlying faculty, of the utmost importance for practical life—the faculty of good sense, common sense, judgment, in a word, the faculty of adapting one's self to circumstances '. Adaptation—' the ability which each individual has of guiding or controlling his adjustments to an ever-changing environment by combining his impressions and organizing his reactions '—this was the Spencerian way of describing how intelligence operates on the simplest levels as on the highest.

Clearly, then, when they spoke of inventing tests for ' intelligence ', Galton, Binet and their followers were using the word as a shorthand label for a highly techincal conception. Thus, when critics like Dr. Heim and Dr. Woolf contend that ' mere environmental conditions—such as differences in schooling or the home background—are more than enough to account for all the observable differences in a mental character like intelligence ', and tell us that ' intelligence involves differences in kind as well as of degree ', they are manifestly using the word in quite another sense. To evade these popular misinterpretations of what was already becoming a colloquial term, Spearman proposed to substitute letters. And in technical discussions I myself have found it helpful to borrow a convention from the physicists and statisticians, and use Greek letters for hypothetical concepts and Roman for empirical assessments, with subscripts to discriminate different sources. Yet even these devices are by no means fool-proof: writers in educational journals gaily talk of a child's ' G ' or ' S ' in the same free and easy fashion as before with no reference whatever to Spearman's own definitions.

But, when all is said, the crucial issue is really this : can we find any tangible evidence to support the conception of an ‘ innate general intellectual factor ’ such as Galton and Binet proposed to measure ? If so, let us keep their name for it; if not, let us bow to our American critics, and expunge the word ‘ intelligence ’ from the scientist's vocabulary.

II. The Evidence

To my mind there can be little question that the trend of recent research fully bears out the earlier distinctions. But, as was to be expected, the advances in psychology, neurology and biology have appreciably modified the fundamental concepts in terms of which those distinctions were expressed. Moreover, the distinctions themselves are now seen to be more like the distinctions between the abstract factors a product than the divisions between the rooms or floors of a material structure. The detailed evidence is discussed in the books and papers cited at the end of this article. Here I need only outline the main converging arguments and the conclusions to which they point.

1. *The Intellectual Factor*. The antithesis between the intellectual aspects of mental activity and the emotional or volitional goes back to Plato and Aristotle. The Greek philosophers relied principally on the observation of outward behaviour. The British school, from Locke and Hume to Ward and Stout, tended to substitute an inward self-observation for the observation of others; and consequently the terminology of the modern psychologist carries a strongly introspective flavour. Such words as ‘ intellectual ’ or ‘ cognitive ’ are apt to suggest that those who are seeking to measure intelligence are concentrating on the narrow field of fully conscious processes — the ‘ eduction of relations and correlates ’ or the ‘ capacity for abstract thought ’—so that ‘ intellectual ’ tends to be contrasted with ‘ practical ’, and thus to exclude it.

From the behaviouristic standpoint the basic distinction is quite simple. To operate any machine—be it a motor-car, a guided missile, a rocket hurtling towards the moon, or the human brain itself—two things are essential: first, power to start and sustain the movement, and secondly, a set of structural mechanisms to guide it towards the goal or target. Plato compares the simplest type of mind to a ship blown by the wind and steered by a rudder. With the fully developed mind, as with the fully-manned warship, the task of steersmanship (*cybernesis*, the noun which gives us our word ‘ government ’) is of necessity more complex, often involving elaborate tactical manoeuvres. But his favourite analogies are drawn from the marshalling of an army or the ‘ government ’ of a well-organized state.

Professor Norbert Wiener, in his book on *The Human Use of Human Beings*, has proposed the adjective ‘ cybernetic ’. I myself would prefer to speak of the ‘ dynamic ’ and the ‘ syntactic ’ aspects of behaviour: (‘ syn-taxis ’ is a Greek military term signifying ‘ organized direction or guidance ’). Here, however, to avoid recondite technicalities I shall follow Professor Vernon's example, and keep to the more familiar name, ‘ intellectual ’, interpreting it in a somewhat Pickwickian fashion. To the teacher who protests that a child's character is quite as important as his intellect I need only say that this aspect has already been covered by Dr. Wall's admirable article on ‘ motivation ’.

2. *The General Factor.* In developing their theory of a *general* factor, Galton and Binet had to rely chiefly on rather speculative conjectures derived from observation, introspection, and the evolutionary doctrines of their day. During the last fifty years, however, far more convincing evidence has become available, derived partly from statistical research and partly from neurological studies.

If we think of the 'factor' primarily as a statistical abstraction, to be used for predicting efficiency in various concrete tasks, there can be no disputing its value: all types of intellectual achievement are correlated with it, though in varying degrees. But, whether or not we give it a label—'intelligence', 'G', or the like—it is admittedly just a hypothetical quantity. Dr. Maddox and others, therefore, contend that any such concept is purely 'metaphysical', by which they apparently mean that it can have no place in any scientific discussion.

However, all such strictures ignore the far more concrete type of evidence drawn from the field of neurology. The earlier psychologists who 'denied the existence of anything that could fairly be called general intelligence' were influenced largely by the doctrine of cerebral localization as preached by the physiologists of the nineteenth century. Of these, the oldest, popularly known as 'phrenologists'—a group of scientific workers whose influence on the views of psychologists and educationists has been far greater than is commonly realized—elaborated the theory that there are definite 'organs' for each of the traditional 'faculties', embedded in the brain like plums in a Christmas pudding. Post-mortem observations on patients who had suffered from localized disease or damage to the brain, microscopical studies of the cell-structure of different parts of the brain, and experiments on animals whose brains were excited artificially by electrical stimuli, seemed fully to vindicate this theory of localized brain centres, though with important modifications as to detail.

However, at the beginning of the present century many neurologists began to suspect that the localization was by no means so clear-cut as had been previously assumed. In America, Lashley, after a long series of experiments on lower animals, put forward a theory of 'mass-action'. But how are we to interpret the nature of this 'mass action' of the brain? Sherrington supplied the answer in a pregnant phrase, 'the integrative function of the nervous system': the business of the brain is to integrate or organize the activities of the individual, and thus to guide and co-ordinate his behaviour so as to meet the needs of the current situation: in a word, its function is what I have called *syntaxis*.

More recently, the anatomists, with their improved microscopical techniques, have demonstrated that the structural differences between one individual's brain and another are quite as striking as those between different brain-areas in the same individual. Sherrington pointed out that almost every tissue—hair, skin, muscle and fat—each has much the same general quality all over the individual's body, with minor regional variations; and he concluded that the nervous tissue could form no exception. Under the microscope the brains of low-grade defectives reveal well-marked peculiarities: as compared with a normal brain, their nerve-cells are scanty and poorly developed; the nerve-fibres that sprout from each cell are few in number and show relatively few ramifications; and both cells and connecting fibres are arranged in a more or less irregular, higglepiggledy fashion, whereas the architecture of the normal brain reveals a neat, systematic and orderly arrangement, like the circuits in the blue-print for a complex piece of apparatus. And, what is still more instructive, these distinctive qualities are visible in every part of the defective's brain. Thus, as Dr. Scholl, one of our most skilful microscopists, has put it, ' intelligence is a characteristic of the whole brain cortex '.

3. *The Genetic Factor.* However, during recent years the main target for attack has been, not so much the assumption of a general factor, but the assumption that differences in this factor are inherited or innate. ' All men ', we are told, ' are created equal—equal not only in political rights but also in mental capacity: the genius is made not born, and the making of him is due to his privileged environment '.

In my view the confusion that still befogs this bewildering controversy has arisen largely because both the believers in mental inheritance and the sceptics still cling to a theory of heredity that is hopelessly out of date. Most of the psychologists who have supported a hereditarian theory have adopted the old Darwinian notion, which supposed that heredity meant ' the principle that like begets like '. Hence they looked more especially for marked similarities between brothers, or sisters, or children and their parents: if the parents differed widely, then, they argued, the child would be a blend of both, much as Aristotle supposed that, if a leopard mated with a camel, the product would obviously be a giraffe. And so, to demonstrate heredity, they relied, as Pearson did, almost exclusively on coefficients of correlation, which measured the resemblances between members of the same family.

During the opening years of the present century, the hypothesis of particulate inheritance, confirmed as it was by a vast mass of experiments on both animals and plants, completely altered the traditional conception. As the modern biologist sees it, the problem of genetics is to explain, not why like begets like, but why in so many instances like begets unlike. Mendel's principles provide the answer. On the basis of a large amount of data, much of it collected by school teachers, I ventured to suggest, in a paper published over forty years ago, that the application of his principles to psychology would help to account for many puzzling features in the transmission of mental characteristics in man.

The family histories of mental defectives and the manner in which individual differences in intelligence were distributed among the general population strongly suggested that here, as with so many physical characteristics, the underlying tendencies were handed on in accordance with the two-fold principles of unifactorial and multifactorial heredity—the wider and rarer deviations being due usually to a single gene, and the rest to the cumulative effect of a large number of genes. Pearson, however, rejected the whole notion of particulate inheritance, except possibly for certain qualitative traits, like colour blindness. The figures for quantitative or ' graded ' characteristics, such as height and general intelligence, could, so he maintained, be accounted for only by the old Darwinian hypothesis of blended inheritance. Psychologists, too, protested that the Mendelian theory seemed to disrupt the total personality into a mosaic of particular traits; and the majority soon decided that, if the champions of the hereditarian theory disagreed so violently over the explanation of the observable facts, then probably both were wrong.

In America most educationists endorsed Bagley's oft-quoted dictum, namely, that ' the Galtonian doctrine of a hereditary aristocracy of talent is wholly out of keeping with the principles of democratic education '. The gifts of the gifted, it was argued, are a social inheritance not a biological; and what parents bequeath to their offspring are not their talents but their privileges. Thus, as Gardner Murphy points out in his book on *Human Potentialities,* the environmentalists seemed virtually to assume that educative influences ' acted on an essentially formless human nature ', or at least on a population whose brains were formed after exactly the same pattern in every case. But, if that were so, then the brain would be an organ that was quite unique in the whole biological world. We may readily agree that the anthropoid brain is imbued with an amazing capacity for learning; but such a capacity, like everything else that is part of the creature's inborn nature, is bound to display marked variations from one individual to another.

Today, however, no geneticist would accept Pearson's criticisms as valid. His own calculations were confined to physical characteristics—height, arm-length, dimensions of the head, and so on; and throughout he adopted an over-simplified method of statistical analysis; nor was any allowance made for assortative mating, which is particularly obvious in the case of mental characteristics and in countries with marital traditions like our own. When his data were worked over afresh by Sir Ronald Fisher with a more adequate mathematical procedure, it appeared that the correlations that Pearson himself had calculated tallied almost exactly with the values deduced from a more up-to-date interpretation of the multifactorial theory. Accordingly, my co-workers and I attempted to apply Fisher's techniques to the analysis of the data supplied by mental tests; and here also we found a surprisingly close agreement.

For those who believe in the old-fashioned theory of blended inheritance, and still more for those who insist on the overriding importance of environment, one of the most paradoxical phenomena is the occurrence of exceptionally bright youngsters in families where the personal, cultural and economic conditions of the home would, one might suppose, have foredoomed them to failure. On the Mendelian hypothesis these anomalies are just what we should expect. It implies that the results of marriage, if not marriages themselves, are a lottery. The random mixing of the parental genes produces as great a variety of recombinations and patterns as we should get by shuffling a pack of cards or shaking the coloured chips in a kaleidoscope.

Moreover, the Mendelian theory enables us to correct one highly misleading assumption made both by the supporters of the hereditarian theory and by its more recent critics. Nearly all of them assume that what is inherited are the characteristics actually observed. Today we recognize that (with rare exceptions, such as differences between blood-groups and possibly in eye-colour) all observable characteristics—the crude measurements of intelligence obtained by tests, just as much as the crude measurements of height obtained with a tape-measure—are the compound products of the interaction between the individual's genetic constitution on the one hand and his environment on the other hand. There are no such things as inheritable traits; there are only inheritable tendencies: and tendencies can never be directly measured, but only indirectly estimated.

The Relative Importance of Nature and Nurture. However, the more cautious critics hesitate to deny outright the possibility that differences in intelligence may be innate: what they question is, whether such differences have any appreciable influence as compared with that of the environment. Dr. Floud and Dr. Halsey, for example, have argued in a recent paper that the tests used in the 11-plus examination are found to favour children from the 'privileged classes' (i.e. the professional and clerical groups): such children do better than the rest simply because their parents were better educated and could, therefore, hand on, by instruction and example, the benefits of their own education. The notion that the parents themselves inherited, and so were able to transmit genetically, a higher measure of general intelligence is in their view wholly mistaken. The most plausible assumption, they maintain, is 'the hypothesis of near randomness in the social distribution of 'intelligence A'' (i.e. of innate intelligence).

How, then, are we to decide between these two alternative views? Whenever we are confronted with two simultaneous influences like heredity and environment, the obvious way to compare their effects is to keep first one and then the other as uniform as possible.

There are numerous orphanages and residential institutions where children of widely different parentage have been brought up from earliest infancy in practically the same environment. When their intelligence is tested, we find that the individual differences are quite as great as they are in the general population; and the case-histories show that the brightest of all are often the illegitimate offspring of fathers from higher social ranks.

To study the effects of environmental differences on children with the same genetic endowment, we can compare assessments for the intelligence of 'identical twins' who have been brought up in different homes. During our work in London schools, Miss Conway and I have been able to trace over forty pairs of this kind: in some cases one twin had been boarded out in a foster home in the slums, while the other had been reared in a cultured and comfortable household. Most of them have been followed up in after life. The detailed figures have been published elsewhere. We found that the correlation between the educational attainments of the twins who have been reared apart was only 0.62, as compared with 0.89 for twins brought up together. On the other hand, the correlation between the assessments for intelligence was as high as 0.88—almost identical with that of twins brought up together.

From all this mass of data, it is possible to deduce, for a given population, an overall figure for the relative influence of nature and nurture. Taking first the crude test-scores, we find that the child's genetic constitution accounts for approximately 77 per cent. of the individual variation—rather less if the tests are merely group tests of a purely verbal type. But no competent psychologist treats the test scores as final. In my own work I have regularly submitted the raw results to the children's teachers; and, wherever discrepancies have emerged, the figures were checked and adjusted by means of further tests and interviews. With these adjusted assessments, genetic constitution appears to account for approximately 87 per cent. of the individual variance.

Professor Vernon reminds us that a number of recent researches claim to have demonstrated 'significant I.Q. rises among feebleminded persons'. But that is merely a reflection on the tests employed or on the way they have been selected or administered; it in no way proves that differences in intelligence are not innate. Nor is the discovery by any means as new as the investigators suppose. When Miss Gaw, Hugh Gordon and I were studying children of gipsies, tramps and canal boat workers, we found that most of them fell below the borderline for feeblemindedness when judged by the Binet tests.

When removed to suitable residential schools, the Binet I.Q's frequently rose by 15 to 20 notches; but the I.Q's obtained with more suitable tests remained virtually unchanged. I know of no reputable psychologist who has ever believed in the alleged ' dogma of the fixity of the I.Q. ', or has ever diagnosed mental deficiency on the basis of an I.Q. alone. From my own experience I should rather say that the use of psychological methods of assessment (including intelligence testing) has more often been responsible for rescuing children from unjust condemnation to a special school or institution than for condemning them to ' vegetate there for the whole of their lives '.

As to children at the other end of the scale, I should readily agree that even the best available methods of selection, when applied on one definite occasion, are " bound to do ' injustice ' to at least 5 per cent.", and that, as Mr. Yates and Mr. Pidgeon put it, " some 10 per cent. are likely to be ' wrongly ' allocated ". But I would prefer to mitigate the wording. Terms like ' unreliability ', ' injustice ', and ' wrong ', tempt the layman to regard them as palpable reasons for abolishing a system which is so fraught with error. But such arguments are rather like inferring that, because the predictions of the doctors who examine us for life assurance are so frequently falsified, therefore the premiums that some of us have to pay are ' unjust '. Any alternative scheme of educational selection would do even graver ' injustice ' not only to the children, but also to the ratepayers and taxpayers who defray the cost of ' wrong ' allocations to the more expensive types of education.

Nor can I agree that, because ' we can never observe Intelligence A ' (innate differences due to genetic constitution), therefore it ' would probably be wiser to forget all about it'. On the contrary, I hold that we should do our utmost to refine our methods of detecting such differences. As I tried to point out in *The Backward Child*, the attempt to train the child who is innately dull is not only a depressing experience for all concerned—teacher, child and parent—but tends to divert effort, time and money from the education of those who are not innately dull but merely backward, and therefore susceptible of marked improvement. On the other hand, children of high inborn ability are so scarce that it is the bounden duty of every education authority to ensure that none of them slips unnoticed through the mesh, but that all are caught at the earliest possible age.

III. Special Abilities and Disabilities

The prolonged and animated controversies that have been waged over the subject of general intelligence have, unfortunately, deflected attention from the still more baffling problem of special abilities. Too often it seems tacitly assumed that the notion of aptitudes for specific subjects was long ago decently interred with other hoary relics of the defunct hypotheses of the faculty psychologist. Its chief opponents, the associationists of the early nineteenth century, like their successors of more recent times (the so-called ' behaviourists ') maintained that all mental progress consists simply in the cumulative effects of ' association ' or (in the modern jargon) ' conditioning ': every child, it was supposed, comes into the world equipped with a miscellaneous assortment of well-defined reflexes; and, as a result of experience or tuition, these become ' associated ' or linked to new stimuli or with new and more effective movements. Spencer's theory turned the whole process inside out; in his view the general capacity for adjustment or adaptation with which every creature is endowed at birth, and which at first is carried out by the body as a whole, tends, as a direct result of growth and maturation, to develop into an increasingly specialized repertory of specialized abilities. There is, as Gesell has recently put it, 'a progressive differentiation within a total action system'.

Most psychologists of the evolutionary school—Bain, Galton, Binet, McDougall—adopted the Spencerian theory. Spearman, on the other hand, at least in his earlier publications, followed the lead of the associationists, and challenged the whole notion of specialized capacities spontaneously maturing. According to his interpretation of the 'factors of the mind', the so-called 'group factors', i.e. the 'special abilities' commonly postulated by the educational psychologists of his day—aptitudes for such subjects as mathematics, literature, art, music, and the like—were merely the outcome of acquired interests or of specialized training at home or at school.

Unfortunately, the neurological evidence is here extremely meagre. A study of the development of cell-branches in the brain indicates that certain areas mature much earlier than others: the centres for movement ripen about ten months after birth, the sensory areas a little later, the areas for speech at about twelve months, and certain association areas (e.g. those in the frontal lobe) not until about four or five years. After that the microscope proves too coarse an instrument to detect any further details. During the school period, therefore, we have to rely for our evidence partly on statistical studies and partly on case studies—especially of children showing marked disabilities in particular subjects.

Systematic surveys usually bring to light occasional cases of specific disability (in reading, for example) among children of normal intelligence long before they have left the infants, department. On the other hand, infantile prodigies (e.g. in numerical calculation and in musical or artistic ability) are reported from time to time by the head teachers themselves. History and biography have already familiarized us with several outstanding instances of precocious genius. Rembrandt painted with an amazing skill at 9; Tennyson covered his slate with childish poems when he was only 8; Mozart published four sonatas at the age of 7; and at the age of $3\frac{1}{2}$ Carl Gauss saved his father 25 thalers by noting a mistake in his addition of the workmen's payroll. But such cases are exceptional. When we turn to the statistical data, it proves exceedingly difficult, before the age of 9 or 10, to demonstrate anything more than the all-important 'general factor'. The influence of special abilities (' group factors', as the statistician likes to call them) is so slight that, unless the surveys cover a vast number of cases, the evidence is altogether swamped by the 'errors' of measurement and sampling.

By the age of 11, however, there are clear signs of a distinction between two broad types: some children seem to belong to what has been termed (in a rather narrow sense) the 'intellectual' or academic type; these, if they are of high intelligence, are obvious candidates for the grammar school; others belong rather to a more practical or mechanical type; and these, provided they possess the requisite intelligence, are better suited to an education with a technical bias. Of course, here as elsewhere, the so-called 'types' are no more than emergent tendencies. Take a random sample of a hundred 11-year-olds, and, on the basis of the most reliable assessments procurable, divide them into three approximately equal subgroups of about 33 or 34 pupils—those who are good at grammar school work, those who are poor at it, and those who are merely average; next reclassify them into those who are good at technical work, those who are poor at it, and those who are merely average: you will then find that, out of the 33 who were good at technical work, as many as 20 will in all probability be good at grammar school work, and only 3 definitely poor at it: in other words, at this early age success in either type of work still depends far less on special aptitude than on an all-round mental efficiency which manifests itself to a varying extent in nearly every type of subject.

Towards puberty the 'intellectual' group tend to subdivide into those of a predominantly verbal type and those with a special facility for numerical work. At the same time there are growing indications of a rich variety of aptitudes provided a sufficiently broad range of tests is employed. Thus, throughout the years of mental growth there appears to be, as Spencer maintained, a continual tendency towards differentiation and specialization.

Of the more specialized disabilities, space allows me to mention only those that bear most directly on the work of the classroom. To begin with, there can be little doubt that several sensory defects run in certain families, e.g. defects of form vision (myopia, hypermetropia, astigmatism, and the like), as well as defects of colour vision which indeed have provided some of the most instructive specimens of Mendelian inheritance in man. Differences in mental imagery likewise seem to be transmitted biologically: some children are almost destitute of visual imagery; others of auditory imagery. But of all such disabilities, the most detrimental is weakness in mechanical memory or retentiveness. In the work of the school, the teacher teaches and the child is expected to learn: hence, pupils who find it hard to retain what they have learnt are apt to drop behindhand in the day-to-day work of the classroom. Defective memory provides one of the commonest explanations why children of comparatively high intelligence at times appear to be educationally subnormal: conversely, in the more mechanical work of the classroom dull children with retentive memories may at times produce unexpectedly good scores in standardized tests, and so present the familiar anomalies noted by those who calculate 'achievement quotients'. As I have shown elsewhere, once these special disabilities have been detected, any appropriate changes made in the teaching methods, the backward child can often be transformed into a perfectly normal scholar.

IV. Practical Implicationa

Let us now recapitulate the main conclusions and see what practical corollaries may be drawn.

1. If, as now seems beyond dispute, each pupil comes into the world with a unique genetic constitution, which in its turn produces wide individual differences in both general and special capacities, then it is plainly imperative that both teachers and local authorities should take full account of such differences in their efforts to provide an education which will (in the words of the Act) be adapted to each child's 'ability and aptitude'. At the primary stage, as we have seen, it is differences in general intelligence that are of paramount influence. When we take a cross-section of the school population at the age of ten, we find that, in a random sample of a thousand, the brightest will have a mental age of 15 and the dullest one of barely 5. By the age of 10 or 11, therefore, some kind of provisional

selection is absolutely essential: in my view it should start much earlier. Provided, as his capacities develop or decline, the child is free to swim from one stream to another, streaming in itself makes for efficiency all round; and, contrary to the prevailing notion, segregation is only needed for subjects that are closely dependent on general intelligence, not for every aspect of youthful life.

2. Since genetic constitution merely determines growth-tendencies, and since the child's observable progress and performance are the resultant of innumerable factors, both internal and external, we could never hope, even with the most precise specification of the newborn infant's equipment of genes, to make infallible predictions about his ultimate level. For the individual pupil the most reliable assessments are unquestionably those compiled by an experienced, observant and well-trained teacher who has watched and recorded his progress from the earliest possible age. But standards differ from one teacher to another, and from school to school. Tests supply a quick and easy way of securing provisional estimates on a scale which is, or should be, the same for all; even so, skill and training are needed to choose, apply and interpret them. Intelligence tests were originally introduced in the L.C.C. schools as a supplementary and standardized procedure for screening those whose backwardness seemed attributable to irremediable defects, and they still have their uses in this field; but they have since proved still more valuable as an accessory instrument for bringing to light what after all are the most precious of our national assets, namely, children of outstanding ability.

3. Owing to variations in environmental conditions and the inevitable fluctuations during growth and maturation, every attempt at allocating pupils must be something of a gamble. But this reflects neither on the efficiency of the tests nor on those who use them; still less is it a reason for denying the importance of inborn differences. It follows that, if we wish to make absolutely certain that the brightest of the bright—say the most brilliant 5 per cent.—will be allotted to a grammar school, then in practice we must cream off a top layer of about 20 per cent. For borderline candidates the accuracy of the assessments can undoubtedly be improved if the results of the group examination are submitted to the teachers or at least compared with the teachers' personal estimates and if cases showing a discrepancy are then subjected to a further check by interview or individual testing. Whether the ultimate gain to the community, and to the borderline children themselves, will be sufficient to offset the additional expense, must be for the education authorities to decide.

4. For those who are unfamiliar with more technical methods of assessment the I.Q. (preferably computed indirectly from percentile rank on a standard scale) furnishes a simple and intelligible unit. But the idea of an H.Q. for I.Q's (if I may so describe it) belongs to the realm of scientific fiction rather than of genuine science. No competent psychologist has ever supposed that an I.Q. can be exactly ascertained, much less that it will remain immutable for each individual child: on the contrary, the earliest researches on mental growth revealed discernible fluctuations from year to year, especially towards the period of puberty. In areas where the threefold classification of secondary schools is operative, an assessment of each child's ability must, of course, be made at 11-plus; but in my view it is equally important to make it at 11-minus—indeed, as soon as possible after the child has entered school. But such assessments should always be checked and

rechecked; and, when at length the child has been allocated to a secondary school, it is desirable that the decision should be subject to review and that a working procedure should be available for transferring or re-transferring pupils during the years that follow.

5. The emergence or maturation of special aptitudes during the pubertal period is an added ground for continuing the study of each individual child throughout his school life. At present they are too often overlooked. Even in the primary school special disabilities may sometimes handicap the brightest; and just because they are specific, it is nearly always possible, as we have seen, by changing the method of teaching, to circumvent any handicap that may thus be imposed. In the secondary school special aptitudes may be almost as important as general intelligence in determining the suitability of a particular type of education or a particular occupational career. As it is, the venerable tradition of the grammar school leads us still to place too much weight on verbal and literary abilities: even our intelligence tests are predominantly verbal. In a technological age—an age of turbo-generators and atomic piles, of automation and electronic control—it is equally essential to discover pupils with aptitudes for technological work.

6. Of the emotional reactions which these administrative changes are liable to engender, I have said little or nothing. In the past children of the same family varied as widely as they do today; but now, with the educational ladder free to all, the bright boy from the poorest home is able to climb, rung by rung, to the highest social rank. However, the different social classes vary far more in their ideals, their traditions, their accent and vocabulary, their code of morals and manners, than they do in actual ability. Hence, the new entrant to the grammar school and university, like the old 'scholarship lad', is apt to feel declassé: he has been uprooted from the group into which he was born, and has no roots in the new. The result is a sense of double frustration, fraught with all kinds of dangers, which Mr. Hoggatt (who has himself been through it all) has graphically described in a book that every teacher should read. Except for a few casual references in studies of maladjustment and delinquency, it is a problem which so far has been almost entirely neglected; and it is one for which much may be done by the schools that now receive the new type of pupil.

In spite of popular prejudice there is no insuperable conflict between equality as a principle of justice and inequality as a fact of genetics. In education equal opportunity means opportunity to make the most of differences that are innate. The ideal is a free and fair chance for each individual, not to rise to the same occupational rank as everyone else, but to develop the peculiar gifts and virtues with which he is endowed—high intelligence if he possesses it, if not, whatever qualities of body, mind, or character are latent within him, irrespective of his intellectual level. In this way and in this way alone will he be able to contribute his utmost to the community, and at the same time live his own life to the full.

7. Finally, as must have been obvious at every stage of our discussion, there are innumerable issues crying out for further research, where the teacher himself can make the most valuable contributions. I have room to touch upon only one or two.

On the practical side we need far more follow-up studies to ascertain where existing methods of selection are faulty and how they can be improved. At present most of the investigations on educational psychology are being carried out by young research-students, often from overseas, with little or no inside experience of classroom conditions or of the teachers' special point of view. The most appropriate investigator (as the L.C.C. long ago realized) would be a psychologist who is himself a member of the education staff and is thus familiar at first hand with the schools in his area and their various problems. Armed with official authority he can plan his experiments more effectively and call for data that are often confidential; while at the same time he will be able to assist or co-ordinate the efforts of teachers in their own investigations.

On the theoretical side the fundamental questions must be left in the main to the psychological laboratory. However, at this transitional stage in our educational organization, the more pressing problems are problems in mental biology—the determination of the genetic basis, not merely of general intelligence, but of special abilities and disabilities, and of peculiarities of temperament and character. These are essentially problems for the field worker. During the earlier years of the century it was the teachers themselves who provided much of the requisite data. They compiled case histories; they traced the pedigrees of families showing conspicuous mental characteristics; and, by their experienced inquiries, they did much to pave the way for what Binet called the ' new psychogenetic approach '. Galton's studies of one-egg twins were mainly based on details obtained in this way; and, if the Mendelian hypothesis be accepted, almost equally instructive material can be secured by comparing the offspring of consanguineous marriages or of parents or grandparents who have married more than once.

Owing to the revolutionary changes that have taken place during the past fifty years, not only in environmental conditions, but in our whole mating system, all these problems require to be worked over afresh in the light of modern genetic theories. Social and geographical barriers have broken down; populations in isolated areas that have been inbred for centuries are now shifting to outbreeding. Changes such as these, as every genealogist is aware, not only furnish some of the most instructive types of source-material, but are also likely to exercise a profound effect on the subsequent history of the group. Surveys and factorial studies, carried out over long periods of time as well as over wide areas, are equally desirable. All these, however, are but a sample of the many ways in which the teacher in the school can aid in the advancement of psychological knowledge. The National Foundation for Educational Research, as we know, is always ready to serve as a co-ordinating or supervising centre.

In a prophetic fantasy which he entitles *The Rise of Meritocracy* Dr. Michael Young has vividly depicted the advantages that might accrue to the community and to civilization if ' educational experts ' could systematically assess the potentialities of every individual as soon as he enters school, and then allot the most suitable training for each one—whether normal, subnormal, or supernormal. Nor does he overlook the drawbacks and possible dangers. And the lesson of it all is that reforms of the type envisaged must be based on an appreciation of the facts, not on political aspirations. Facts, however, can only be ascertained by impartial inquiry. Hence the two crying needs of the moment are a more scientific attitude towards the issues involved and a more determined organization of scientific research.

References

Ballard, P. B., *The New Examiner*. London: University of London Press, 1925.

Board of Education, *Report of the Consultative Committee on Psychological Tests of Educable Capacity*. H.M. Stationery Office, 1925.

Board of Education, *Report of the Consultative Committee on the Primary School*. H.M. Stationery Office, 1931.

Burt, C., " The Inheritance of Mental Characteristics." *Eugen. Rev.*, IV, 1-33, 1912.

Burt, C., *The Distribution and Relations of Educational Abilities*. London: P. S. King, 1917.

Burt, C., " The Evidence for the Concept of Intelligence." *Brit. J. Educ. Psychol.*, XXV, 158-177, 1955.

Burt, C., " The Differentiation of Intellectual Ability." *Brit. J. Educ. Psychol.*, XXIV, 76-90, 1954.

Burt, C. and Howard, M., " The Multifactorial Theory of Inheritance and its Application to Intelligence." *Brit. J. Statist. Psychol.*, IX, 95-131, 1956.

Conway, J., " The Inheritance of Intelligence and its Social Implications." *Brit. J. Statist. Psychol.*, XI, 130-140, 1958.

Fisher, R. A., " The Correlation between Relations on the Supposition of Mendelian Inheritance." *Trans. Roy. Soc. Edin.*, LII, 399-434, 1918.

Fisher, R. A., " The Causes of Human Variability." *Eugen. Rev.*, X, 213-220, 1919.

Floud, J. E., Halsey, A. H., and Martin, F. M., *Social Class and Educational Opportunity*. Heinemann, 1956.

Galton, F., *Hereditary Genius*. London: Macmillan, 1869.

Galton, F., *Inquiries into Human Faculty*. London: Dent, 1883.

Halsey, A. H., " Genetics, Social Structure and Intelligence." *Brit. J. Sociol.*, IX, 15-28, 1958.

Hoggart, R., *The Uses of Literacy*. London: Chatto and Windus, 1957.

McDougall, W., *Psychology: The Study of Behaviour*. Home University Library. London: Williams and Norgate, 1912.

Murphy, Gardner, *Human Potentialities*. New York: Basic Books, Inc., 1958.

Pearson, K., " On a Generalized Theory of Alternative Inheritance with Special Reference to Mendelian Laws." *Phil. Trans.*, CCIII, 53-87, 1903.

Piaget, J., *The Psychology of Intelligence*. London: Routledge & Kegan Paul, 1950.

Sholl, D., *The Organization of the Cerebral Cortex*. London: Methuen, 1957.

Spearman, C., *The Abilities of Man*. London: Macmillan, 1927.

Terman, L. M. and Oden, M. H., *The Gifted Child Grows Up: Twenty-five Years' Follow-up of a Superior Group*. London: Oxford University Press, 1947.

Vernon, P. E., " A New Look at Intelligence Testing." *Educational Research*, I, 3-12, 1958.

Weiner, N., *The Human Use of Human Beings: Cybernetics and Society*. London: Eyre and Spottiswoode, 1950.

Young, M., *The Rise of Meritocracy*. London: Thames and Hudson, 1958.

The Teaching of Mathematics—1

THE DEVELOPMENT OF NUMBER CONCEPTS IN YOUNG CHILDREN

by J. B. Biggs

Introduction

BASIC to any attempt to improve the teaching of mathematics is a knowledge of the way in which young children form number concepts, and how the teaching of arithmetic and number experiences in the primary school affect this growth. Although arithmetic itself has been the subject of research for many years, much of the reported material in the extensive literature tends to be fragmentary and disjointed. It seems, therefore, that an attempt should be made at the outset to sketch a psychological scheme of interpretation on which a systematic body of knowledge, more directly useful to teachers, may be built.

Buswell, in a review of the previous 20 years of research in arithmetic, has a similar sort of thing in mind when he writes that too much time has been spent in dealing with incidental problems—analyses of textbooks and curricula, comparisons of specific techniques of teaching given topics, and so forth—and far too little on the thinking being done by children as they learn. When this second question can be answered, i.e. when a systematic psychology of arithmetic emerges, the former incidental problems will be well on the way towards a solution. The emphasis, in other words, is to be changed from *what* children know to *how* they come to know. Largely because of the difficulty in undertaking such research, there is, with the important exception of Piaget's work, little to guide us on this second and crucial point. The Scottish Council, for example, reports a survey of the extent of number knowledge amongst children entering infant school. This is certainly useful information; but much more useful would be an account of how these number concepts develop in the first place, and how the teacher can fit her method of teaching to this natural process of concept development.

The more important questions seem then to fall into two groups: (*a*) how does the child acquire understanding of number concepts? and (*b*) what are the implications which a knowledge of this growth has for our teaching methods, the curriculum content, the age at which any concept or process might be taught, and so on?

It can readily be seen how this second group of questions is dependent on the first. Appropriateness of method and the placement of topics in the curriculum, for instance, are to be determined not so much by the logical demands of the subject to be taught, as by the psychological demands or limits of the child. It is an evident waste of time to teach a child decimals, say, if his mental equipment is at present inadequate to allow him to understand them. It could be argued that the concept .1 is logically the same as 10, in relation to 1; but psychologically this may be beyond the grasp of children who can in fact understand—functionally at any rate—the latter relation.

The actual contents of the curriculum may be governed in some part by educational aims and philosophies, and by such external factors as eventual social and vocational requirements; nevertheless, the real content must inevitably be limited by the understanding of the pupil, i.e. by what he can ' take '. In the past, this has been in part determined by a so-called common sense psychology; by such assumptions, for example, that drill develops a disciplined mind; or that mere continued exposure to bodies of knowledge will eventually result in some at least being assimilated. Modern educationists are somewhat more humble in their psychological claims.

It is worthwhile, then, to review such theoretical contributions as we have to the psychology of number. The most thoroughgoing and potentially fruitful source of relevant theory is to be found in the work of the Swiss psychologist, Piaget ; and it is on this that much of what follows is based. His researches were motivated in the first place by a somewhat philosophical aim—to find a genetic basis for a theory of knowledge—and are only incidentally, as it were, a contribution to education. In spite of this, his work, and that of his students and more recent followers, is of paramount importance to arithmetic teachers, especially those concerned with primary school pupils.

Piaget's Contribution to the Psychology of Thinking

Piaget has been interested in mental growth in general and particularly in the way in which children develop concepts of space, time, causality, reality and number. His method of study was to interview children individually, giving them simple but highly ingenious tests to complete, and questioning them afterwards on the reasons for their responses. On the basis of many years of such work he distinguishes several stages of development paralleled in most of the variables he investigated.

In *The Psychology of Intelligence* he defines the following general levels of intellectual growth, giving *approximate* age levels in each case:

I. Non-operational (Birth—2 years). This he also calls the *Sensory-Motor* stage. The infant is solely concerned with objects *qua* objects—if a ball is temporally hidden from the baby, for him it ceases to exist. Having no power to construct a symbol for the ball—i.e. lacking the concept ' ball '—the infant is completely tied to the here-and-now world of palpable objects.

II. Pre-operational. (1) *Transductive Period* (2—4 years). As the abstraction from action to thought is made, memory becomes possible. The object, with the development of language, is symbolised, however crudely, by a thought process which can be retained in the mind. Even so, there is no distinction as yet between the general and the particular which is an essential part of the formation of true concepts. The three-year-old uses what Piaget calls a ' pre-concept '; for example, he does not know whether he eats the same grapefruit for breakfast each morning or a different one since he lacks the general concept of a class of object—grapefruit—to which a single fruit belongs.

(2) *Intuitive Period* (4—7 years). Here the child's thought processes are governed entirely by global perception—if an object *looks* bigger than another, it *is* bigger, regardless of the circumstances. Thus, another child who is standing close by is pronounced to be taller than a man some distance

away. His conceptual apparatus is as yet unable to regulate his sense impressions: the imaginings and fantasies of the six-year-old are thus not so imaginative or fantastic as appears to amused—or horrified—adults.

III. Operational. (1) *Concrete* (7—11 years). Two important principles become available at this stage. The first is the reversibility of thought. This is exemplified in the test situation where the child is asked to lay out a glass for each of several bottles on the table—a task even very young children can do without much difficulty—and he will see that if the line of glasses exactly corresponds in length to the line of bottles, there is the same number of glasses as bottles. However, if the glasses are pushed together, the five-year-old would say that there are now too many bottles (because the line of bottles is longer) and would 'equalise' the situation by putting out more glasses. He thus makes his judgment on what he immediately sees before him, without any reference to his preceding actions. His thought processes, in other words, are irreversible. Taking a more mathematical example, the proposition $2 + 6 = 8$ necessarily implies its reverse, $8 - 6 = 2$. A five-year-old may be 'taught' to perform both these sums with accuracy (as indeed may a parrot), but he will fail to appreciate the fact that the second is the opposite of, and is implied in, the first. When he can appreciate this, it means psychologically that he is able to range his thought backwards and forwards over a complete experience.

This principle of reversibility permits him to make the next important step which is to discover the constant elements in such operations. Thus, when he repeatedly finds that $2 + 6 = 8$ he will conclude that 8 is always 8, whether it is 8 rabbits, motor-cars, pussies or just plain units; similarly for 2 and 6. In fact, there is some evidence to show that an insistence on 'enlivening' number work which is apparent in many infant schools may actually impede the young child's progress in understanding the essential constancy of number meanings. Stern, as we shall see later, adds strong support for this view in her book, *Children Discover Arithmetic*.

The principle of conservation, as this new development is called, is already well recognised in the physical sciences and here is extended to the conceptual world of thought. The notion of set, or of any material or conceptual closed system, requires that in any transformation of its components, or their interrelationships, the total set will remain unchanged. The concept '4', for example, still remains constant whether it is regarded as $3 + 1$, or as $2 + 2$. Piaget tested this notion of conservation with his celebrated liquid and glass experiments, and found three stages in the development of the concept. At first (absence of conservation) children would respond that the quantity of liquid increases or decreases according to the size or number of its containers. Thus, if a glass of lemonade is poured into a taller but thinner glass, the five-year-old will almost invariably respond that there is more in the second glass (because the level is higher) or that there is less (because the glass is narrower). Next, at about $5\frac{1}{2}$—6 years, there is an intermediate stage when the child will judge correctly that there has been no change in volume as long as the second glass is not *too* tall or *too* thin—i.e. ' common-sense ' triumphs over crude perceptual evidence if they are not too much at variance from each other. If the difference is too marked, however, the child will revert to the perceptual criterion.

At the third stage, constancy becomes immediately apparent—indeed, seven-year-olds tend to get a trifle impatient at being asked what is, at this stage, all too obvious. By substituting beads for water, Piaget obtained similar evidence, and drew similar conclusions, for the conservation of discrete (unitary) quantities. All forms of constancy are not, however, attained at this age. Inhelder points out that at seven years the child can arrive at conservation of matter (by moulding a ball of plasticene into a long strip), but it is not until ten that he understands conservation of mass or weight, and eleven for conservation of volume.

The child at this level, then, is capable of a limited form of reasoning, as long as it is tied to his own concrete experience. He is able to generalise and to form valid concepts but only so far as they are founded in and known immediately through experience.

(2) *Formal* (11 years and over). At early adolescence, ability to perform abstract operations becomes apparent. Children can, for instance, pay attention to the form of an argument or situation and ignore the actual content. They can construct a rational *as-if* world; set up hypotheses and deduce the logical consequences. This liberation from the bonds of the concrete is achieved by the final stage of concept formation whereby mental structures are ' decentred ' from perception.

The differences between this and the previous stages can be seen by referring to Inhelder's work where children were asked to discover elementary physical laws by manipulating simple apparatus. At the pre-operational stage, children played with the apparatus in random fashion, quite without any notion of an underlying principle. Subjects at the concrete operational level varied one condition at a time, and could note a cause-and-effect relation, but in particular instances only. Adolescents, however, carefully and systematically controlled the known variables and thus induced, in true scientific manner, the explanatory principle behind the observed phenomena.

This final feat of abstraction is impossible at an earlier age or lower level of mental development. It is worth noting in passing that only at secondary school level can we thus be reasonably sure that topics of a formal nature—ratio and proportion, algebra, scientific method, theoretical geometry, logic, etc.—are potentially fully understandable to children. On the other hand, as Peel warns us, this does not mean that all our thought after the age of 11 years is at the formal operational level. We all, he says, tend to aspire to the condition of operational thinking, but if the tasks are very new and very difficult, then even the adult will find himself having to move through some of the preliminary stages.

The Child's Conception of Number

The earliest age at which it would be at all possible to observe behaviour remotely relevant to the formation of number concepts would be at the pre-operational stage. Verbalisation is a most important pre-condition, but is not, in itself, sufficient—as is evidenced by the infant who can happily count (correctly) up to 10 or more without the slightest clue as to how to *apply* these pleasant and diverting sound-patterns. Objective evidence for this is found in Descoeudres' work. She discovered that, whereas at three years most children could count up to 5, yet they were unable to perform matching operations (where some understanding of the meaning of number is required) beyond 3 objects. Five-year-olds, on the average, could count up to 9, but were able to attach number names to only 3 or 4 objects.

To recapitulate briefly : in Piaget's view, concept formation is the result of 'action experience' building up mental structures which then serve as a reference system for the interpretation of direct sensory—or 'physical'—experience. These mental structures permit reasoning as such; basically, they are abstractions or generalisations from concrete experiences which impose order upon direct perceptual events. We have seen, for example, in the glass and bottles experiments that plain global perception is apt to be most misleading. Prerequisite to all thought are the basic structures of reversability and conservation. We shall now look at some of those additional ones which are specific to mathematics.

Logically, our number system is composed of two basic processes—cardination and ordination. Any number, that is to say, may be regarded as possessing a definite characteristic or class property, e.g. 'four-ness'; and also as indicating a place in a series, e.g. the fourth member of the series, preceding 5 and succeeding 3. It is worth pointing out, as Churchill has also noted, that emphasis in infant schools is all too frequently only on the cardinal properties of number which themselves cannot be understood without the 'action experience' of ordination: on the other hand, systems of teaching which are based only on counting are equally remiss. Mere verbal counting, as Piaget insists and Descoeudres has demonstrated, does not adequately account for all the ordinal properties although it is vital for both cardinal and ordinal understanding.

One of the important features of cardination is the ability to classify. In the experiment with bottles and glasses it will be remembered that, despite appearances, the notion of invariance of class 'bottle' in relation to class 'glass' was of prime importance. This goes a stage further when the child is asked to classify objects and then to combine them into a new class—an operation which is the basis of ordinary addition. To demonstrate this, Piaget gave his children some brown beads and a much smaller quantity of white ones; both groups of beads were made of wood, i.e. they all fell under the concept 'wooden bead'. The child at stage one correctly maintained that there were more brown ones than white ones, and agreeing that both brown and white are wooden, would then go on to make the surprising claim that a necklace made out of brown beads would be longer than one made out of *wooden* beads 'because there are more brown ones than wooden ones'. Not until the age of seven are Piaget's children able to hold two concepts in mind simultaneously and thus to grasp the logical (and mathematical) truth that the part is never greater than the whole. In other words, until children do see this, they are incapable of *understanding* addition or subtraction. The point at issue here, though, is the demonstrated relation between classifying and the recognition that a class, or combination of classes, has a definite number of elements.

There are also three stages in the development of ordination. At first the child is only capable of global ' greater than ' or ' less than ' judgments. He is thus able to construct a simple series, say of sticks of varying lengths, but such a series is strictly limited in size. At the second stage he is able to seriate well enough, but when the series is disarranged internally he is unable to correct it by reference to the number of elements preceding a given place in it. His understanding, that is, is qualitative, not truly quantitative. Finally, cardination is co-ordinated into the process so that each element in the series has a definite value; that this value is determined by the number of preceding elements; and that the unit value is constant, i.e. the difference between 2 and 3, say, is the same as that between any other adjacent members of the series, e.g. 51 and 52.

These stages may be illustrated by a study reported by Churchill. One of her tests involved the use of Tillich bricks. These are blocks ranging in size from the smallest unit (cube) to ten times the length of the cube, the unit sizes being painted in colour along the sides of the rectangular blocks. It was thus possible to build a series either by reference to the length of a block in relation to the others (qualitative seriation), or by counting the units marked off on the larger blocks—in effect, determining its cardinal value (numerical seriation). Churchill found that very few five-year-olds could at first complete a ten-step series and those who could (2 out of 16) were unable to answer questions relating to the positional value of the elements in the series. Later, however, stage three type responses were obtained in which the children spontaneously counted the squares on each brick to determine its position in the series; i.e. the cardinal and ordinal aspects had become integrated. Further, these children could justify the position of a given item by saying that ' it was a four ' (i.e. four bands or units were painted on the side of the block) and therefore occupied *fourth* position. In contrast, typical stage one and two responses were: ' Because it's littler ' or simply, ' I don't know '.

These, then, are some of the basic processes which Piaget investigated. It would be pointless following him through *The Child's Conception of Number* in tortuous detail. The major mechanisms have been outlined above and there is ample evidence here for us to draw some vital, if not too comforting, conclusions. First, however, a note on interpretation. Much of the data comes, especially in the earlier stages, from Piaget's own three children; hardly, as Berlyne puts it, ' the children of Average Man! ' Further, one may feel some qualms in generalising from the responses of a small and biassed sample of Swiss children to the situation of the English Infant and Junior school. However, considering the intensive and clinical nature of his research, and that his main point is to indicate the sequential and structured nature of mental growth, these objections are more specious than real. Most of the comparative evidence we have—and there is an ever-increasing amount—differs from Piaget's conclusions in detail only.

The University of Birmingham, for instance, has been carrying out a good deal of work to test and generalise Piaget's conclusions. Amongst these studies, one reported by Carpenter relates to age levels at which changes take place. She administered some of Piaget's tests to a group of children, aged from 5 to 9 years, and ranging, in each age group, from dull through average to bright. She found the same sequential nature of intellectual development but the ages at which the stages occurred varied markedly from those reported by Piaget. She suggests that this may partly be due to the degree to which the child is familiar with the test material and partly to the amount of interest the test material itself arouses in the child (the children did better on one test when the material consisted of model trains instead of plain balls). Probably her most interesting finding, however, was that there was far greater agreement between conceptual development and *mental* age than with chronological age *per se*. Unfortunately, Piaget related development to the latter only.

Lunzer confirmed Carpenter's findings with a larger sample. He concludes that, as expected, their work is in agreement with the findings of Piaget but that stage level is better determined by mental than by chronological age. Further, it is suggested that where the context of a problem is familiar, a subject can tackle such a problem at a higher level than where the problem is set in an unfamiliar context. The teacher, he goes on to say, might by frequency of presentation of a type of problem thus facilitate the reasoning process in his pupils—" in so doing the teacher is deliberately educating children, rather than waiting for development to take its ' natural ' course."

A recent study conducted by Churchill confirms this view that the Piagetian stages are not purely maturational; i.e. that environmental influences can be important. She criticizes Piaget for not paying sufficient attention to the influence of cultural and educational factors on the developmental process—for " if children of five or six years have as little understanding of the meaning of numbers as Piaget suggests, then the little sum cards they fill up every day in many schools must hold a very different meaning from the one intended by the teacher, and they could be more profitably occupied in other activities."

Sixteen five-year-old children were used in the enquiry. They were first tested by a series of the Piaget type and the group divided into two so that mean initial scores were the same. The experimental group was given a programme of play situations involving grouping, seriating, matching and ordering so that the children " might learn something about the nature of number operations and relations." The control group simply carried on with their usual Infant school routine. The testing was repeated after a month. It was found that the number understanding of the experimental group had improved significantly over that of the control group, as is indicated by the much larger number of correct test responses and the higher proportion of stage three responses.

These results in the main support Piaget's own conclusions but there are two practical modifications: the sequence of stages is apparently not so invariant as at first supposed (some children, for instance, seemed to by-pass stage two); and secondly, the process may be hastened and facilitated by appropriate experiences—it is not essential, in other words, to wait for the gradual maturing of inherited potentialities.

On the other hand, there has been some violent disagreement over his work. Much of this is purely theoretical—e.g. Bunt's criticism concerning the philosophical use of logical concepts—but there is one study which appears to refute Piaget's results quite unequivocally. Estes reports that four (only) of Piaget's tests were administered to 52 Kentucky children ranging in age from four to six, and she comes to the conclusion that if children could count correctly, rearrangement of objects did not alter their reported number, even with the four-year-olds. However, she omits to include a qualitative analysis of the children's replies, and since her apparent reduction of number understanding to counting is contrary to most (non-Piagetian) evidence, her results cannot be taken as being conclusive.

But for our purposes the real worth—or otherwise—of Piaget's psychology* is best seen in its paedogogical implications. Three things seem to follow from a consideration of the child at any of the Piagetian stages. Having found that a child is unable, say, to construct a one-one correspondence, we can say: (i) he is not yet *ready* to go on to more complicated mathematical operations such as adding, because (ii) they would have no *meaning* for him, and therefore it is likely that (iii) he will do poorly even in mechanical arithmetic—even though he may be trained to perform at a relatively high degree of speed and accuracy by dint of sheer drilling. This last, incidentally, is a good reason why mechanical tests of arithmetic are, by themselves, inadequate measures of ability to 'do' arithmetic. Fortunately, there is quite a considerable body of independent research evidence on each of these three points.

Some Educational Implications

1. *Arithmetic Readiness*

Piaget makes it quite clear that his children at ages four to six years do not possess the understanding of number sometimes attributed to them by their teachers. Verbal fluency with number words and ability to do rote counting do not allow us to conclude that the child knows what he is talking about. We have seen from the results of the brown and white beads test, that prior to stage three the child is largely incapable of appreciating the essential features of addition and subtraction. But this does not mean that one must wait until a chronological, or even a mental, age of seven-plus has been attained before teaching addition. Churchill's work demonstrates that, with appropriate teaching methods and number experiences, the requisite level of conceptual understanding can be developed as early as five to six years. The implication is rather that *until* this understanding is developed, at whatever age, it is a waste of time teaching such operations as 7 + 2 = 9. Churchill would go even further than this. Premature arithmetic teaching, she says, may well be doing something more harmful than wasting time: it may set up barriers to the proper understanding of number when the child *is* ready. She suggests that this may be one of the causes of "the strong dislike and even fear that many intelligent adults have of mathematics."

Williams considers that the two main problems relating to early number readiness are that there is a lack of any readiness tests, i.e. tests of conceptual attainment, and that there is a need for a practical classroom programme to foster number thinking especially in those children who have poorly developed number concepts. He believes that the best approach to these problems is to be found in Piaget's psychology. Accordingly, he constructed two test batteries. The first consisted of representative tests from Piaget's own work; the second was a simple attainment test to relate computational ability to an apparent level of conceptual development. In comparing the results of these two tests,

---

*For a rather more detailed exposition of *The Child's Conception of Number* and a discussion of the educational implications, the reader is referred to a booklet published by the National Froebel Foundation: *Some Aspects of Piaget's work.*

he found that some children were able to compute accurately enough, but lacked any appreciation of the numbers concepts involved. Such children, on further investigation, were unable to make any but the simplest generalisations. It must remain, Williams concludes, a matter of conjecture as to what proportion of these subsequently become unhappy and insecure in relation to number activities.

He suggests a readiness programme, consisting of the basic process of ordination and cardination, which should assist the child in acquiring operational number concepts. When these are attained, and not before, the child is ready to proceed to more formal work. It is probably worth noting that in many other countries children do not begin school until six years of age, and in Sweden for example, where the age of commencement is as late as seven, teachers are forbidden to take mechanical arithmetic during the seven-plus year and may use only concrete apparatus. It is perhaps not so surprising, then, that the standards of such countries around the secondary transfer level are not noticeably inferior to our own—in spite of our apparent start of two years in learning arithmetic.

In an unpublished investigation conducted by the N.F.E.R., it was found that amongst thirty-eight primary schools studied, there was no significant correlation between amount of formal work attempted in the infant school and later success in the junior school as measured by mechanical and problem tests. In other words, those children who had covered a good deal of formal work (table-learning, manipulation of large sums, etc.) in the earliest stages evinced no tendency to be superior on later testings to children who covered a relatively less amount in the infant school.

Other relevant non-Piagetian evidence at this point may be found in the following studies. During the 1930's, the "Committee of Seven," under the chairmanship of Washburne, was set up to investigate the grade-placement of topics in the arithmetic curriculum in the United States. With teaching methods and time allocation controlled, each topic was introduced at three or four consecutive grades in a population of over 30,000 children in sixteen States. A battery of intelligence tests and 'foundation' tests (to ascertain mastery of previous knowledge required for a given topic) was given. The topics were then treated in class. Finally, a battery of attainment tests was given to measure mastery and retention of the new topic. The minimum age at which 80 per cent. (or sometimes 75 per cent.) of the children had mastered each item was determined. Several of the results are of interest. It was found, firstly, that when tested at Grade VI, the best performances came from those who began formal arithmetic at Grade I (age 6 +); next came those who began at Grade II (age 7 +); and finally, those who commenced at Grade III (age 8 +). However, and this is the interesting point, if a child did not begin formal arithmetic until Grade II or III and instead was given *informal* number activities and number games, his later performance was very much superior to the child who began formal arithmetic in Grade I. In any case, in terms of economic success, i.e. achievement in relation to the time spent in teaching, later placement was indicated. The Committee recommended that the addition facts with sums of 10 and under, and easy subtraction facts, should not be attempted before the mental age of seven to eight years; that is, well into the Junior level in our own schools! Up to this level, the curriculum should consist of "informal experience and activities to give children real concepts of number and space relations, without any systematic drills."

The validity of Washburne's findings can be challenged on many grounds. Curr, for instance, replicated the study as far as was possible in Scottish schools, but using only the teaching of division, and found that mental age *per se* was a misleading guide to placement. He came to the conclusion that a better prediction of future success could be made from the foundation test itself (mastery of prerequisite skills), or even better, from intelligence test results. Not having the same power to control teaching methods as Washburne, Curr found that there was a wide divergence in " optimum mental ages " for topic placement and that this seemed to be largely related to the techniques employed by individual teachers. Some teachers, that is, can teach the division process successfully at an earlier age than others—and there seems no valid reason to suppose that this would not be true of other branches of the subject. It would thus be wrong to accept the Washburne findings literally. Rigid adherence to them may hamper the efforts of the efficient teacher; and it would mask the effects of general improvements in teaching method.

There is, however, no dispute about Washburne's theoretical viewpoint; that " there is a point in a child's mental growth before which it is not effective to teach a given process in arithmetic, and after which that process can be taught reasonably effectively." The disagreement is in the precise determination of that point. It is clear that, although practical considerations such as school organisation and teaching techniques will bear on it, this is mainly a psychological problem. Placement and readiness studies will be only partially satisfactory if they are not rooted in child development and it is here that we find the relevance of Piaget's genetic approach.

Finally, it may be mentioned that on the matter of postponing formal number work, Brueckner, in a critical review of Washburne's and other relevant studies, states:

> " Strong evidence, therefore, leads one to conclude that formal work on number processes can be delayed until at least the second, even to the third grade, but that, if that is to be done, steps should be taken to develop in the first one or two grades a series of well-organized activities and units in social applications of number through which the child will be brought into contact with number in meaningful social situations in such a way as to insure more adequate readiness for the more formal work to follow."

2. *Teaching Methods: " Meaning Theory "*

The effect of psychological theorising on education can be clearly seen in the popularity from time to time of different approaches to the teaching of arithmetic. Basically, this has come down to mechanical versus meaningful teaching. The research results are also useful in interpreting the recommendations of Brueckner and Washburne concerning the postponement of formal work in the early years.

Up to the 1930's treatises on the teaching of arithmetic stressed an approach derived from the connectionistic, stimulus-response laws of behaviourist psychology. This consisted essentially of learning " bonds " between numbers. Hence, the theory ran, the more frequently the pupil practised drill, the more firmly " fixed " the number facts became in his mind. The technique was first to propound a general rule which was then learned by application to *n* examples.

About this time, however, the Gestalt school of psychology became popular and very different implications for teaching method resulted. The emphasis was upon " structured wholes " and the inter-relationship of parts to the whole. Applied to arithmetic, this implies that the understanding of number relations and of the total structure of arithmetic should be aimed at: an approach obviously more suited to problem solving than to drill—indeed, drill was deliberately neglected as being mechanistic and therefore inimical to the development of number meanings.

An approach sometimes confused with the meaning method is " social " arithmetic. This demands that the topics in arithmetic be limited to those which are found in adult usage—indeed, the extremists restrict this to child usage only. Thus, one would not give a pupil an example involving hundreds of pounds as it would be most unlikely that he would ever have handled such a sum. Studies of the use made by adults of arithmetic (e.g. Wilson and Dalrymple), have shown that in the area of fractions, 90 per cent. of business dealings and ordinary usage are covered by halves, thirds and quarters; the remaining 10 per cent. being covered by eighths, decimals and possibly twelfths. These and similar findings have resulted in the purging from textbooks of examples involving meaningless (to the child) LCM's of thirteenths, sevenths, etc., and have, indeed, led to much more realistic and meaningful examples and problems. However, it is obvious, in spite of the somewhat extravagant claims made for social arithmetic, that it is primarily concerned with the curriculum, although the problem-types suggested by social usage have been eagerly fastened upon by the meaning theorists. There is a definite link between the child's usage of a topic and the meaning that topic will have for him (see, for example, Brideoake and Groves' Birmingham experiment on infant number work) but this particular aspect of number meaning will be discussed later.

Although the controversy between the advocates of mechanical and meaningful teaching has been raging for about 25 years, as recently as 1957 Miller could point out that the relative superiority of the meaning method was as yet undetermined: and that " little research has been initiated to justify the great emphasis on the approach."

Nevertheless, one of the most impressive studies on the meaningful approach is that reported by Brownell and Moser. They obtained the co-operation of teachers in presenting four different ways of teaching subtraction, each teacher concentrating on one method. Using decomposition* and equal additions*, each was presented rationally (where the meaning of the place system was explained) and mechanically (where proficiency in learning a sequence of acts was emphasised). 1,400 Grade III children (approx. age 9) were used in the experiment. The results were evaluated not only by speed and accuracy but also according to degree of understanding, length of retention, transfer (i.e. the ability to transfer a set of rules or an understanding from one situation to another, where appropriate) and the value of crutches. Among their conclusions were the following: (i) decomposition taught rationally was the best method; (ii) equal additions was hard to rationalise in terms the pupils understood, but equal additions

---

*Take the example: 134
69
——
65
——

In the decomposition method, one would say " 9 from 14 leaves 5 ; put down 5 and carry 1 ; 13 becomes 12 ; 6 from 12 leaves 6."

According to the equal-additions method, " 9 from 14 leaves 5 ; put down 5 and carry 1 ; 6 and 1 are 7 ; 7 from 13 leaves 6."

N.B. The wording of these illustrations is more in line with the mechanical approach.

taught mechanically obtained very speedy results—much more so than mechanical decomposition, or equal-additions-rational; (iii) Crutches were found to be helpful in decomposition especially, and could be discarded by the pupils without ill-effect when they were directed to do so by the teacher; (iv) generally, the rational approach led to much better retention over a period and to greater ability to transfer to new situations; while the mechanical procedure tended to be faster but less accurate—hence less economical ultimately. The success of the rational method will depend to a large extent on the degree to which any given topic can in fact be rationalised successfully in terms that children can grasp.

Anderson, using division, found that, overall, there was no appreciable difference between the two methods. On closer analysis of his results, however, it was found that for so-called " over-achievers " (i.e. those whose performance in arithmetic was better than might be expected from intelligence test results) drill was more effective; but with " under-achievers " (who would more likely be found in remedial classes under present selection methods) teaching which stressed generalisation and understanding was to be preferred.

Howard found that when fractions were taught by any of three methods (drill; meaning; meaning followed by drill), there was no *immediate* difference, but that after a period, retention tests showed a high loss for the first method, less for the second, and no loss at all for the third.

Both Thiele and Swenson found that children trained by the meaning method in learning the 100 addition facts were significantly more able to transfer this knowledge to arithmetic problems than those who had learned them by rote.

Miller, with a group of 600 children, went a stage further than those studies and investigated not one process, but all the processes involved in the elementary school curriculum, and in addition, employed a " Meaning Test " designed to evaluate the child's understanding of number concepts. His results are not entirely unequivocal, but he indicates that the meaning method is superior for dealing with fractions, decimals, percentages and " difficult concepts ", and also produces higher retention, especially with pupils of average and high ability. The rule method, on the other hand, appears to be more effective for immediate teaching of measurement, and generally for children of low intelligence.

On the other hand, Schonell recommends that children should attain automatic accuracy in the addition and subtraction facts—indeed, he advocates that addition tables should be learned by rote as should be the multiplication tables. Meddleton gave one group of children large amounts of systematically graded drill sums, so that all number combinations up to 100 were practised, with extra practice on combinations known to be difficult, and found that the children so drilled performed at a considerably higher level in both mechanical and problem arithmetic than those who had been given textbook sums to practise haphazardly. Thus, if used intelligently, even large amounts of drill need not be at all deleterious. But neither Schonell nor Meddleton would advocate drilling to the point of " automatic accuracy " as the means of *introducing* a topic.

Another finding of the N.F.E.R. study mentioned earlier was that where teachers in the fourth year of the junior school adopted a meaningful approach to their work, as indicated by a positive reply to a question on whether they explained the “ invert and multiply ” rule in division of fractions or merely applied it mechanically, they obtained significantly better results in mechanical and problem arithmetic generally.

These results of a number of investigations, though not absolutely concordant, may lead us to certain valid conclusions:

1. If speed in a highly specialised task is the aim of teaching, the drill or mechanical method is likely to be superior: if, however, the children are required to transfer their knowledge and skill to new situations, the meaning method is much to be preferred.
2. Mechanical training may be more effective if the skill is required immediately after the training period, but if the skill is meant to be at all enduring, meaningful teaching is almost a *sine qua non.*
3. When dealing with more complicated material and more difficult concepts, the meaning method is essential. However, with children of low intelligence, the meaning method may be too difficult to apply: in which case (if the skill in question *must* be mastered) one would have to fall back on the rule method.
4. There is some evidence to the effect that children *like* to be taught meaningfully —although it must not be forgotten that the rote learning of material and the calculation of large numbers of drill sums can be source of pleasure to many children. The pleasure, however, soon palls if the child is continuously at a loss to understand what he is doing—even if he is getting his sums correct.

It seems, generally, that the most effective method is a combination of both—the topic should be introduced in a meaningful way, after which a reasonable amount of practice, or drill, is desirable to gain familiarity with the new skill or principle.

For the infant teacher, however, the question still remains: what is the best method of *introducing* arithmetic to children? The evidence seems to point to the meaning method, but we need to know *how* to explain something as basic as early number work, and guidance is necessary on the rather vexed question of the use to be made of concrete materials.

Prior to Piaget's work, there was little known about the logical properties of numbers *in relation to* the corresponding psychological development of number concepts. The problem of meaning could thus only be approached in a limited way. Either one analysed the logical aspect (*cf.* the “ unit-skills ” approach, as exemplified in Brueckner's early work) or, in the manner of the Washburne Committee, one obtained statistical surveys of the levels of skills at which children were capable of performing given operations, under given conditions of teaching. Neither approach offered much insight into the sequential and developmental aspects of number understanding and hence their applicability to teaching method, especially when we think of individual differences, is limited.

A sensible adaptation of Piaget's principles makes it possible, in the first and second years of school life, to provide the “ rich number experiences ” indubitably indicated by other research findings. If a proportion at least of these are of a “ social ” nature, so much the better. It must be stressed, however, that these experiences should cover the areas of number operations which Piaget analysed, i.e. conservation-producing activities, matching, seriating, corresponding, additive- and multiplicative-compositions and so on. It is useless to prescribe such activity for a stated class or year level. The goal is " stage three understanding " whenever this may be attained. After this, the explanatory (meaningful) approach to specific topics—addition, division, etc.—is indicated. The proportion of drill will depend upon such things as the nature of the task, the intelligence of the pupil, the aim of introducing the topic and the desired length of retention.

Piaget explains the development of mental structure as involving, first, the activity of the child in his environment. This implies that children need much concrete experience in manipulating objects. Thus, school work which demands excessive abstraction before the child is ready will only puzzle and frustrate him. On the other hand, as this is only a *stage* in normal development, " spoon-feeding " by restricting the child to simple apparatus for too long may inhibit the development of the formal operational stage.

It is worth mentioning also that Van Engen, in an analysis of meaning in arithmetic, comes to the conclusion that action-experience is an essential prerequisite to the development of meaning. He derives the conclusion from a consideration of semantics and current meaning theory: meaning is acquired only when the symbol of an act is tied operationally, i.e. by " doing ", to that act or event. This theory springs almost entirely from logic; nevertheless, its educational implications, in this respect, closely parallel those following from Piaget's findings.

3. *The Place of Concrete Materials*

It was noted earlier (p. 19) that, contrary to some "progressive" views, indiscriminate enlivening of number work by means of concrete materials may actually be harmful to the development of number concepts. In the introductory chapters of her book, Stern subjects the whole question of the value of concrete materials to a searching analysis. Her point is that it is just as " abstract ", from the child's point of view, to present him with a picture of 4 bunnies and 3 bunnies (= ? bunnies), as to teach him by rote that $4 + 3 = 7$. What have *bunnies* to do with *units?* The important thing that the child has to learn is the relation between 4 and 3—which " bunny-learning " will tell him no more readily than rote learning. Churchill quotes a classic example of this sort of thing. She was visiting a classroom where the children were playing a fishing game. The fish were made of blue and pink card, each fish bearing a number. When a blue fish and a pink fish had been caught, the total score was to be added. Brian, aged five, caught a 3 and a 5, and when asked by the teacher how many he had caught, he replied, " two. " This was not the " right " answer, as the teacher tried to explain.

In a recent survey conducted by the N.F.E.R., some incidental information about teachers' attitudes to concrete materials was obtained. There were great divergences in viewpoint—some maintained that concrete materials should be used only when essential; others that they baffle the child, that he should be introduced to the abstract nature of mathematics as soon as possible; while yet others regarded them as quite essential to arithmetic. Obviously, not all the teachers were thinking of quite the same thing, but there is, even so, a valid difference of opinion about their worth. This difference could well be due to a failure to make the distinction noted above between material which is concrete, but essentially non-mathematical, and that which is a " concretising " of the properties of the number system. In this same study, a comparison was made between the school which employed a great deal of concrete material and those which used somewhat less. There was no significant difference in scores when the children were tested on mechanical problems arithmetic at each age level in the junior school. This result points to the conclusion that the use of concrete materials *per se* is not in itself advantageous: rather the use made of them seems to be the important thing. Ives reports that successful teaching of backward and emotionally disturbed children was most readily accomplished by material (Stern) which made concrete the structure of number: even very dull children could acquire certain ordinal and cardinal concepts which were simply not attainable by children taught with the more usual matchstick-counter types of material. On the other hand, Williams has noted that younger and more backward children may tend to use such material merely as an aid to computation, in as rigid and mechanical a manner as the counting of the traditional cowrie shells.

Further evidence that concrete materials may be abused is suggested in a study reported by Murray. She investigated the use of concrete materials in four Scottish infant schools, and found that when the children were tested during the first and second years in the junior school, those who had been introduced to arithmetic *via* the extensive use of concrete materials performed at a considerably *lower* level in mechanical, problem and mental arithmetic, than those who had been trained with little concrete material, the effects of intelligence having been partialled out.

On psychological grounds it is clear that the aim of number education is to procure understanding of the meaning of numbers and then to apply this to numerical processes, in mechanical and, later, problem form. Past practice in the infant school has possibly reversed this. By constant repetition of number processes the child, it is hoped, will gain understanding of number: a hope, it may be added, often realised in the case of very bright children. We should expect, then, that materials providing experiences which further numerical understanding should be of great value: those which simply put drill in a presumably more pleasant form are of more questionable help.

Of the former kind, there are first those suggested by the activities which are prerequisite to number understanding—matching, seriating, etc. Churchill suggests many such activities in her thesis, as also does Williams in the article cited earlier. It may well be, too, that the success of " social " projects is not so much because they are adult situations in miniature, although they would appeal to children for this reason, but because shopping, for example, gives experiences of one-one correspondence (one penny for one sweet); both weighing and measuring emphasise conservation, the constancy of the unit and cardination proper; and so on. Other valuable activities and materials can easily be devised by the teacher. The important thing is to provide equal opportunities for the often neglected ordinal aspects of number, as well as the purely cardinal, to develop.

It is interesting to note that Gunderson gave seven-year-old children, who had not been specifically taught multiplication and division, some problems involving both these processes. She found that when they fully understood the processes of addition and subtraction, they could, with the aid of concrete and semi-concrete materials (counters and diagrams), generalise to the more complicated processes of multiplication and division without any help from the teacher. Two things should be noted here: (i) if the children had been taught addition and subtraction by a mechanical method of following rules, they would not have been able to solve the test problems, and (ii) it is unlikely that they all could have performed this feat of generalisation without being tied to reality, as it were, by the materials provided—although a few of the brightest children were able to solve the problems mentally.

There are more elaborate materials which attempt to represent the actual *structure* of the number system. The basic idea behind these is that the child, by manipulating rods or blocks in a systematic and controlled way, will discover for himself (under a certain amount of guidance, of course) the meanings of number relations and number concepts generally. The two most popular of such methods are probably those devised respectively by Stern and Cuisenaire, in each of which both counting and measuring—ordination and cardination—are systematically employed. The major purpose of this type of material is to develop initial understanding of relatively simple but basic concepts, although it is also claimed that their use at later stages in the junior school provides the conceptual foundations for such topics as directed numbers, manipulation of unknowns, and so forth, in secondary mathematics.

In Piagetian terms, one could say, firstly, that by manipulating the material himself, the child is gaining action-experience out of which operational concepts may develop; and secondly, this mode of presentation is in keeping with his conceptual level since the pre-adolescent is not capable of true formal abstractions. Thus, his perception of number operations symbolically structured by blocks helps him to grasp the particular concepts so represented.*

There has, however, been very little research reported to date where the effectiveness of such material has been unequivocally demonstrated: the N.F.E.R. is aware of this *lacuna* and is initiating a large-scale investigation to determine what actual, practical value they might have. Such work as is reported, however, seems to bear out theoretical expectations. J. Miller, for instance, taught the 42 children in her experimental group throughout Grade II by the Stern method, while a control class of 45 children, with the same amount of teaching (½ hour each day) and using the same syllabus, were taught by traditional methods. By the end of the year the experimental group had surpassed the controls in all portions of the curriculum, including purely factual knowledge, and especially where understanding of the number system is involved, e.g. place value and problem work.

These findings are highly suggestive, but the scope of the study is too restricted to make valid generalisations permissible. Karatzinas and Renshaw sent a questionnaire to forty teachers who had started at Primary I level to use the Cuisenaire material in Edinburgh schools. While some were reluctant to commit themselves after only one year's experience with the material, the majority considered that they were achieving comparatively better results, that skill in the four rules was more quickly and easily developed, and that the children found the rods more attractive than more usual types of counter apparatus. In addition to this survey, a group of forty children who had had eighteen months of experience with the material were matched with a control group of thirty-eight children who had been taught by traditional methods. When a simple fractions tests and sections of the Schnonell Diagnostic Arithmetic Test were administered, it was found that in the sub-tests completed by both groups there were no significant differences in scores, i.e. the experimental group had gained a no less effective mastery than the control group. However, the control group had not made sufficient progress to be able to attempt the multiplication, division and fraction items, whereas the children taught by the Cuisennaire method could solve these items quite easily. As with the Miller's study, these results are certainly promising but stand in need of confirmation from studies performed on a wider scale.

---

*It is no mere verbal accident that we talk of "seeing" a problem as of seeing an object.

In conclusion, it should be added that when it comes to the practical question which faces the teacher—what methods or techniques ought I to employ?—the answer is by no means as clear as statistical results, even when backed up by respectable theory, might indicate. It is a truism to say that teaching is large a matter of personalities and personal relationships. There is little doubt that a poor teacher, using a very " good " method, i.e. one which has proved successful in a controlled experiment, will obtain worse results than an excellent teacher using " inferior " techniques. Further, the demands of a " good " method may not be suited to a particular temperament; some teachers may well do better, and feel happier, with a " poorer " method. This, for instance, may go a long way towards explaining the results of Murray's study. There has been very little research into the question of teacher personality, very largely because it is a most difficult matter to investigate—in this country, especially.

The best research can do is to point out the effects that certain methods and techniques have had when used under controlled conditions. Where the outcomes coincide with the teacher's aims, and where these conditions approximate to the teacher's own, he has reason to give the findings his most serious consideration.

Conclusions

Consideration of the findings of Piaget and other research workers has led to several conclusions which seem to be of importance to the teacher.

1. The immediate goal of early number teaching, prior to the introduction of formal work, is seen to be " stage three understanding ".
2. This may be attained by means of appropriate number games and experiences involving mathematically sound concrete materials, i.e. those which take both ordination and cardination into account, as early as 5 or 6 years.
3. It would appear to be advantageous to substitute such a programme of number experiences for much of the formal work which is currently performed in infant schools.
4. Psychological theory suggests, and empirical evidence confirms, that meaningful and explanatory teaching in the junior school produces results more consistent with the basic aims of number teaching. Mechanical drill, however, may be valuable when applied systematically to material which has already been introduced meaningfully.

References

This is a list of the material consulted and summarised in this paper. Items providing an initial bibliography for further reading are marked with one asterisk (*); more difficult and specialised items with two asterisks (**).

Anderson, G. L., " Quantitative thinking as developed under connectionist and field theories of learning." *Learning Theory in School Situations*, 40-73. University of Minnesota Studies in Education, No. 2. Minneapolis, 1949.

Berlyne, D. E., " Recent developments in Piaget's work." *Brit. J. Ed. Psych.*, 27, 1-12, February, 1957.

*Brideoake, E. and Groves, I. D., *Arithmetic in Action*. University of London Press, 1954.

Brownell, W. A. and Moser, H. E., *Meaningful versus Mechanical Learning: A Study in Grade II Subtraction*. Duke University Research Studies in Education, No. 6. Durham, N.C., 1941.

Brueckner, L. J., " The development of ability in arithmetic." In *Child Development and the Curriculum*, Thirty-eighth Year Book, N.S.S.E., Part I, 275-298. University of Chicago Press, 1939.

Bunt, L. N. H., *The Development of Ideas of Number and Quantity according to Piaget*. Groningen, Wolter, 1951.

Buswell, G. T., " Needed research on arithmetic." In *The Teaching of Arithmetic*, Fiftieth Year Book, N.S.S.E., Part II, 282-297. University of Chicago Press, 1951.

Carpenter, T. E., " A pilot study for a quantitative investigation of Jean Piaget's original work on concept formation." *Ed. Rev.*, Institute of Education, University of Birmingham, 7, 142-149, February, 1955.

**Churchill, E. M., *Early Number Concepts: an experimental study of the growth of numerical ideas in five-year-old children*. Unpublished M.A. Thesis, University of Leeds, 1956.

**Churchill, E. M., " The number concepts of the young child." *Researches and Studies*, University of Leeds Institute of Education, No. 17, 34-49; No. 18, 28-46; 1958.

*Cuisenaire, G. and Gattegno, C., *Numbers in Colour*. London: Heinemann, 1954.

Curr, W., " Placement of topics in arithmetic." *Studies in Arithmetic*, Vol. II, 183-218. Scottish Council for Research in Education. U.L.P., 1941.

Descoeudres, A., *Le développement de l'enfant deux à sept ans*, 233-282. Paris: Delachaux et Niestlé S. A., 1921.

Estes, B. W., " Some mathematical and logical concepts in children." *J. Genetic Psych.*, 88, 219-223, 1956.

Gunderson, A. G., " Thought patterns of young children in learning multiplication and division." *Elementary School Journal*, 55, 453-461, 1955.

Howard, C. F., " Three methods of teaching arithmetic." *California Journal Ed. Res.*, 1, 3-7, 1950.

Inhelder, B., " Les attitudes expérimentales de l'enfant et de l'adolescent." *Bull. de Psych.*, 7, 272-282, 1954.

**Inhelder, B. and Piaget, J., *The Growth of Logical Thinking*. London: Routledge and Kegan Paul, 1958.

Ives, L. J., Interim finding obtained in an ongoing experiment concerning the use of the Stern apparatus. Unpublished memorandum to the present author.

*Karatzinas, D. and Renshaw, T., " Teachers' views of the Cuisenaire method." *The Scottish Educational Journal*. 19th and 26th September and 3rd October, 1958.

Lunzer, E. A., " A pilot study for a quantitative investigation of Jean Piaget's original work on concept formation; a footnote." *Ed. Rev.*, Institute of Education, University of Birmingham, 8, 193-200, June, 1956.

Meddleton, I. G., " An experimental investigation into the systematic teaching of number combinations and arithmetic." *Brit. J. Ed. Psych.*, 26, 117-127, June, 1956.

Miller, G. H., " How effective is the meaning method? " *The Arithmetic Teacher*, 4, 45-49, March, 1957.

Miller, M. J., *Investigating the structural arithmetic program of Catherine Stern*. Unpublished Master's Thesis, Boston University School of Education, Massachusetts, 1950.

Murray, M. S., *An investigation into the merits of much use of concrete materials in arithmetic teaching in the infant department*. Unpublished B. Ed. Thesis, University of Edinburgh, 1941.

Peel, E. A., *Learning: in School and Society*. Unpublished report of Teachers' Conference, 1-26. University of Birmingham Institute of Education, 1957.

**Piaget, J., *The Psychology of Intelligence*. London: Routledge and Kegan Paul, 1950.

**Piaget, J., *The Child's Conception of Number*. London: Routledge and Kegan Paul, 1952.

*Schonell, F. J., *Practice in Basic Arithmetic*. Edinburgh: Oliver and Boyd, 1954.

Scottish Council for Research in Education, *The Early Development of Number Concepts*. London: U.L.P., 1942.

*Stern, C., *Children Discover Arithmetic*. New York: Harper and Brothers, 1949.

Swenson, E. J., " Organisation and generalisation as factors in learning transfer and retroactive inhibition." *Learning Theory in School Situations*. University of Minnesota Studies in Education, No. 2. Minneapolis, 1949.

Thiele, C. L., *The contribution of generalisation to the learning of the addition facts*. Contribution to Education, No. 763, Teachers College, Columbia University, New York, 1938.

*Van Engen, H., " An analysis of meaning in arithmetic." *Elem. Schl. J.*, 49, 321-329, February, 1949; 395-400, March, 1949.

Washburne, C. W., " The work of the Committee of Seven on grade-placement in arithmetic." In *Child Development and the Curriculum*, Thirty-eighth Year Book, N.S.S.E., Part I, 299-324. University of Chicago Press, 1939.

*Williams, A. A., " Number Readiness." *Ed. Rev.*, Institute of Education, University of Birmingham, 11, 31-46. November, 1958.

Wilson, G. M. and Dalrymple, C. O., " Useful Fractions." *J. Ed. Res.*, 30, 341-347, 1937.

CLASS SIZE AS A VARIABLE IN THE TEACHING SITUATION

by C. M. Fleming

(Reader in Education, University of London, Institute of Education)

Introductory

THIS is a matter of perennial interest. The teaching situation (as thought of in connection with formal schooling) involves a relationship between adults and children. Is the size of the group significant? Should every pupil have the undivided attention of a teacher? Can equal benefits or greater benefits come from the assembling of several children (or many children) under the direct guidance of one adult?

The question invites consideration of educational purposes. Its answering must obviously be determined also by the resources available—whether financial, material or personal. It is, therefore, not surprising that treatment of the topic has passed through a number of phases, varying to some extent with economic conditions in different centuries and in different countries and closely related to the prevailing undertones of contemporary educational discussion. Like many other areas of educational research, the field was first entered by the philosophers. They directed attention to some of the ideals of effective learning and offered certain suggestions. Echoes of their arguments still linger in our ears; but a clear progression can be traced in the last few decades from such theoretic discussion, through early efforts by psychologists who used the methods of what was then called ' experimental education ', to controlled comparisons with the use of objective tests, to more skilled analysis of test results and, more recently, to fuller recognition of the complexity of the teaching situation and the part played in it by attitudes and relationships.

I. Philosophic Discussion

Lively disagreements are the chief characteristic of the theoretical approach. These range from Comenius' plea for the assembling of large numbers in schools to Locke's distrust of every form of public schooling.

" I maintain," said Comenius, " that it is not only possible for one teacher to teach several hundred scholars at once, but that it is also essential; since for both the teachers and their pupils it is by far the most advantageous system. The larger the number of pupils that he sees before him the greater the interest the teacher will take in his work . . .; and the keener the teacher himself, the greater the enthusiasm that his pupils will display. To the scholars, in the same way, the presence of a number of companions will be productive not only of utility but also of enjoyment (for it gives pleasure to all to have companions in their labours); since they will mutually stimulate and assist one another."

As if in direct reply to this came the rejoinder of John Locke: "Let the master's industry and skill be never so great, it is impossible he should have fifty or an hundred scholars under his eye any longer than they are in school together; nor can it be expected that he should instruct them successfully in anything but their books. . . . Till you can find a school wherein it is possible for the master to look after the manners of his scholars . . . forming their minds to virtue . . . you have a strange value for words when, preferring the languages of the ancient Greeks and Romans to that which made them such brave men, you think it worth while to hazard your son's innocence and virtue for a little Greek and Latin."

Similar diversity of viewpoint is to be found in the 1945 challenge by W. C. Bagley to Carl Seashore's plea for the economic use of inspired teachers and for the encouragement of self-help through larger classes than were then customary. Arguments in favour of large numbers of pupils under the care of one teacher are no longer supported (as they were in the pages of Comenius) by the plea that such classes are the only conceivable means by which the benefits of education can be made available to all pupils irrespective of their prosperity or sex. It is now, in most parts of the world, an accepted tradition that for primary and secondary education groups will be considerably below one hundred in size and that teachers must be found and buildings provided in numbers sufficient to make classes of much less than fifty the rule rather than the exception.*

The connotation of the word ' large ' has thus insensibly altered; and the discussion has widened to include as arguments in support of large classes not only the economical use of buildings, and the reduction in the number of teachers required, but the socialising and democratising opportunities provided by larger groups and the encouragement which they can give to independent study; while on the other hand the plea for small classes now presents the claim that a reduction in size will be followed not only by higher achievement, but by more regular attendance and better behaviour on the part of pupils along with improved morale among teachers.

Opinions of Teachers and Pupils

Much of this discussion comes from parents, administrators, or philosophic observers of the educational scene. Not dissimilar diversity of viewpoint is, however, reported in the smaller number of published statements which proffer a record of the collected comments of teachers and pupils. Teachers on the whole appear to believe that large classes are exhausting, a cause of frustration, and a reason for failure in basic subjects such as reading or arithmetic. In contradiction to this, some find larger classes more interesting and stimulating; and pupils have recorded the view that a greater variety of opinions can be obtained in large classes and that larger classes foster more self-reliance.

When specific statements are invited there is some tendency for preferences to be shown for the size of class to which pupils have become accustomed and in which teachers have experienced success. Over the decades there has been, for example, a gradual reduction in the numbers described as ' large ' and this has run roughly parallel with the lowering of authorised and of actual enrolments in schools in Western Europe and in the

---

*The study of the reasons for this change in viewpoint fall in the field of the history of education and no attempt is made to deal with them here.

United States of America. In 1929, for example, a consensus of opinion among teachers (the majority of whom had classes of over forty) was expressed in favour of groups of not more than thirty-five to forty, while in 1939 the number considered desirable was of the order of thirty. By 1949 (when the median was nearer to thirty) many teachers agreed that though groups of eleven to thirteen were too small for efficient instruction, and groups of thirty-two or thirty-three were too large, the ideal might usefully be set at about five-and-twenty. (Not quite irrelevant to this is a careful study of sub-groups of different sizes within a larger class in which it was reported that greater enthusiasm for work with small groups was found among the more inexperienced teachers and among students highly concerned with their own fears and anxieties while non-authoritarian instructors of fuller experience set more value on the intellectual stimulation of the larger groups.)

Connotation of the Term ' Class '

A further complication in all such argumentation has been the varying connotation of the term ' class '. This may be defined as the number of pupils who regularly meet in one ' administrative and instructional unit—usually under the direct guidance of a single teacher '; and in most reports the word undoubtedly carries this meaning. Where surveys have been extensive there is, however, no reason to doubt the inclusion of certain single-teacher schools and it is rarely indicated whether the word ' class ' is then taken to refer to the number of pupils under the exclusive supervision of one adult or to the number of pupils of a particular grade or age-group within the larger unit of the school.

The significance of the term also varies with accepted practice in larger schools at different stages in the educational process. At the primary stage it is quite customary for a ' class ' to be a group of pupils under the charge of one adult for all or most of the subjects of the infant or the junior school, while at the secondary stage some degree of specialising is usual; and at the tertiary or University stage teachers do not accept responsibility for work in more than one field.

This difference in itself has probably some bearing on the attitude of teachers towards groups of varying sizes; and it may account for the fact that University teachers have tended to be less vocal on the issue of class size than their colleagues whose work lies with younger children. Differences in the degree of specialisation have also tended to be followed by greater variations as between the burdens of one teacher and those of another at the secondary and the tertiary stages where selection of subjects and courses is commonly followed by marked differences in the sizes of the groups electing to take subjects such as music, art, classics, or science in contrast to groups following prescribed courses in certain other subjects.

Teaching Load

As a development of these discussions there has come a gradual realisation of the complexity of the administrative issues involved in both planning and furnishing buildings in which teaching is to take place. In a classroom of a given size a group of little children requires less space than does the same number of adolescents or of adults. Their output in written exercises is also less voluminous and appears to be less exacting from the point of view of the time required for its marking. For reasons such as these it has often been assumed that classes at the primary stage may justifiably be larger than classes at the secondary stage.

At the same time certain changes have taken place in interpretation of the teaching process, and at all levels there is now more doubt as to the effectiveness of the teacher who, as an instructor or lecturer attempts all the time to command the attention of all beholders. " When I do ope my lips let no dog bark," is no longer supposed to be a typical expression of a teacher's point of view and the dreams entertained in the 1890's by Search, Rice and others as to the possibility of more active co-operation in learning on the part of pupils have been followed by the administrative proposals of the 1930's and the new buildings of the late 1940's and the 1950's.

Associated with this is the fact that many arguments about the significance of class-size have themselves been contributory to a more realistic awareness of the nature of a teacher's ' load '. It is now, for example, frequently admitted that this includes not merely the number of pupils in a class but all the activities which are related either directly or indirectly to a teacher's professional duties, responsibilities and interests. Account is, therefore, beginning to be taken of such differences as the character of the subject to be taught, the number of class-periods of instruction per week, the extent of duplication in the preparation of lessons for different classes, the length of teaching periods, the time spent on extra-curricular activities, the extent of overcrowding in classrooms, the amount of secretarial work required, the adequacy of the teaching materials provided, and the total number of pupils with whom the teacher has to make contact. All these are variables in the teaching situation from an administrative point of view; and to an increasing extent argumentation is turning to them in contrast to an earlier preoccupation with the issue of class size.

II. Research Studies

A somewhat similar increase in subtlety of interpretation—though in a different direction—may be seen in the course taken by educational research over the same decades. The nature of the conclusions reached by fully reported research studies on the matter of class size is, however, in marked contrast to the trends of opinion characteristic of the more theoretic approach. Most investigators started from the presupposition that size of class would prove a significant determinant of the degree of success achieved by a teacher. There have been many investigations; but, with few exceptions, it has had to be reported that, under typical conditions, class-size in itself appears to be an unimportant factor. The benefits of small classes, though commonly taken for granted by theorists, are as yet largely undemonstrated in the pages of accredited research reports.

This conclusion has been reached at every level from infant-room to University lecture-theatre. It has been formulated in relation to many subjects; and it is supported both by test results and by assessments of various types.

Available evidence is difficult to interpret—in the first place because of the diversity of opinion as to what constitutes a ' large ' group and what should be described as a ' small ' one. Methods of reporting research findings also range from incidental mention in the analysis of responses obtained for other purposes to deliberately-designed experiments with the use of the most modern statistical techniques in the treatment of results.

Early Psychological Studies

The first investigation which can be said to fall within the field of educational research seems to have been one carried out by Rice in 1902 as part of his enquiry into the reasons for the wide differences among pupils to which attention was then beginning to be given. Using specially-devised group tests of arithmetic he sought to relate variations in the responses of some six thousand children in grades four to eight (ages approximately 10 to 14) to differences in the size of the class, the number of minutes given to the teaching of arithmetic each day, the method employed, the age of the pupils and the nature of the home environment. The conclusion reached in each case was that, contrary to expectation, there was a lack of direct relationship between success and each of the concomitants examined.

It is a matter for regret that in the case of class-size, Rice did not publish details of the scores obtained; but his conclusion is quite unequivocal. " The number of pupils per class was larger in the highest six schools than it was in the schools of City VI (next to the lowest in average achievement) and the classes were exceptionally small in the school that stands at the lower end."

In a repetition of the same experiment in 1903, using a test of language instead of tests of arithmetic, Rice reached the same conclusion. " It has always been supposed that the size of the class must necessarily exert a powerful influence on the results. But investigation showed that there was no relation between the size of class and the results, that some of the best work had been done in the largest classes, and some of the poorest in the smallest classes." These findings attracted considerable attention and the topic of size of class took its place among other contemporary studies of the reasons for scholastic success and failure.

The most productive period from the point of view of research appears to have been the years after the first world war—the 1920's and the early 1930's. Butsch, for example, writing in 1934 mentioned two hundred and five references to one hundred and eight studies cited by Irwin in the years from 1900 to 1932, while Otto and von Borgersrode, in 1941, reported a total of two hundred and sixty-seven studies which it was then possible to trace. Not all of these documents are now accessible and in some instances it is probable that the differences between groups of different sizes were so small as to fall within the limits of those which might have occurred by chance.

Illustrative of a careful early study is one carried out by Dawe in nine schools in Minnesota with four hundred and thirty-three kindergarten pupils arranged in seventeen groups varying in size from fourteen to forty-six. Their recollection of a story which had been read to them was tested by questions on the following day; but no demonstrable differences were discovered either in relation to the size of the group or to the sitting position of the pupils in relation to the teacher. A further study with the same number and size of groups but involving group discussion showed, however, a difference in favour of the smaller groups in the percentage of pupils taking part in the discussion, the number of children who spoke and the number of remarks per child. There was no attempt to estimate the significance of these differences; but the children in the larger groups appeared to have had fewer opportunities for conversation although the amount they learned in a more formal lesson had not been affected.

A contemporary enquiry by Miller into efficiency of instruction in High School Physics showed a similar absence of any appreciable effect on learning among much older pupils with groups varying in size from three to thirty-one. (In this, again, no information was given from which the significance of any differences could be calculated.)

More adequately controlled experiments with the use of objective tests are reported in surveys of research studies in the early 1930's; but in many cases in these also no attempt was made to report findings except in terms of percentages or averages—without fuller statistical treatment of the results.

A summary of the recorded conclusions of the less accessible of these earlier studies may be of some interest.

Table I

A SUMMARY OF STUDIES ON CLASS SIZE

Numbers in Classes		Stage	Author and Date	Differences		Subject
Large	Small			Significance reported	Significance not reported	
44	20 34	Secondary School	Tope, 1924		Medium & small groups slightly better	English
104	36	College	Taylor, 1929		No difference	Psychology
44	23	High School	McGuire & Hudelson, 1930		Mediocre pupils did better in large class	Plane Geometry
70	35	High School	Davis & Goldizen, 1930		Small advantage to large class.	History
51	20 or 21	High School	Smith, 1930		Small class superior in letter writing. Large in reading, etc.	English
80	20		Brooks, 1930	No difference		English Composition
55	30	High School	Bloomfield, 1930	No difference		History
105	27 and 45	College	Beauchamp, 1930		No difference	Education
336 (*largest*)	8 (*smallest*)	College	Hudelson, 1930, 1932	Large superior in 50 per cent. of experiments. Small superior in 5 per cent. of experiments.		140 experiments with groups of differing size.
150	36 and 40	College	Remmers, 1931		Large slightly better	Psychology

Table I—A SUMMARY OF STUDIES ON CLASS SIZE—*continued.*

Numbers in Classes		Stage	Author and Date	Differences		Subject
Large	Small			Significance reported	Significance not reported	
40	20	Grades I to IV	Whitney & Willey 1932		6 per cent. in favour of small class	
42 38	27 20	High School	Holy, 1932	No difference		English
100	35	College	Raths, 1933		No difference	Biology
150	40	Grade VI	Stoddard, 1934		Large gained more	Science and Music with Sound Pictures
105	22 25	College	Hand and Smith, 1934	No difference		Business Training
39	11	High School	Dalthorp, 1934		No difference	English Composition
Over 35	Less than 25	High School	Baker, 1936	Knowledge of pupils slightly better in small		English, etc.
More than 30	Less than 25	Primary	Newell, 1943	Teachers of small classes more inventive and more willing to try new practices.		

More Skilled Analysis of Test Results

Illustrative of the difficulty of arranging controlled experiments is a study by Jensen in which one hundred and eighty-five boys and girls were arranged in groups of from ten to seventeen (the small classes) and from forty to forty-eight (the large classes). The pupils were paired on age, on an intelligence test and on a test in algebra. They were then taught for a semester by teachers who each used the methods they themselves preferred. Retesting at the end of that time, while it showed similar mean gains in all groups, indicated also more significant differences in the smaller groups and a small mean superiority of those in the small classes over their peers in the large classes. These differences were, however, so small as to fall within the limits of those which might have occurred by chance; and it is worthy of note that the pupils knew the purpose of the testing and that rivalry between large and small classes was keen throughout. Something more than merely the effect of size obviously formed part of the experimental situation on this occasion.

A recent study of a different type is that of Spitzer in which, using nearly two thousand pupils in grade III and over seventeen hundred pupils in grade VI, a series of comparisons (with calculation of the significance of the differences of means) was made between one hundred and five classes whose average size was about twenty-two (the small classes) and forty-three classes of thirty pupils or more (the large classes). The tests employed were the Iowa standardised tests of reading, language skills and arithmetic skills and the classes tested were deliberately selected from both large and small cities in such a way that it could not be said that the larger classes were predominantly in the larger places which might perhaps have attracted the more capable teachers and whose level in equipment, buildings and supervision might have been higher.

After extensive analysis of test results it had to be reported in this series also that no significant differences were found as between the two sizes of groups although slightly lower scores were made by the smaller groups in language skills and slightly higher scores in arithmetic and reading at approximately age nine along with slightly higher scores in arithmetic by the larger groups at about the age of twelve.

Findings Incidental to Other Enquiries

Not dissimilar findings are recorded as incidental to another study in Maxwell's report on the 1947 Scottish Mental Survey in which a test of general mental ability had been given to all pupils aged eleven years. The size of the class was in this case defined as the " ratio of pupils on the roll to the number of teachers "; and in general it was found that while children with the highest average test scores (38.09) appeared in the very small schools with fewer than twenty pupils and in the very large ones with more than nine hundred pupils (37.97), the highest average test score, when classes of different sizes were considered, was that of pupils in classes of ten to nineteen children (38.11). Next in order to this came the averages 37.21, 36.46, 36.32 and 36.12 in classes of sizes twenty to twenty-nine, one to nine, thirty to thirty-nine, and forty to forty-nine respectively. The significance of these differences was not reported; but it was added that although a few schools with a ratio of more than fifty pupils per teacher were recorded for the cities there was no indication that these pupils were significantly better or worse than the others in test scores. There was also no apparent relation between test score and the number of schools previously attended.

This evidence as to the superior scores made by pupils in the very small schools in the more remote rural areas may be compared with the findings of another report from the Scottish Council for Research in Education on the testing of pupils of average age twelve and three months in some two thousand four hundred schools (Mowat). In this study the highest average quotients in tests of Intelligence, of English and of Arithmetic were obtained either by schools presenting fewer than ten pupils for admission to secondary schools (those in Arithmetic) or by schools presenting more than thirty pupils (those in Intelligence and English). The differences between these and the scores of pupils in schools of medium size were large enough to be significant. Information is not here available as to the exact size of class in each case; but in the circumstances of the testing it was known that the schools presenting fewer than ten candidates were for the most part the small schools in rural areas while those presenting more than thirty were schools with large classes in urban areas.

More relevant to the question of the direct differences associated with size of class was material recently obtained* by the National Foundation for Educational Research in England and Wales from a sample of more than four thousand pupils aged eight years. Using tests of mechanical reading ability (fifty words) and a picture 'intelligence' test, a comparison was made of results in classes of less than twenty with those in classes of more than twenty pupils on the grounds of the higher average scores made in the Reading test by pupils in classes of two to nineteen (53·33) as compared with those in classes of twenty to twenty-nine (51.70), thirty to thirty-nine (50·33), forty to forty-nine (51.28) and fifty to fifty-eight (50.75). Analysis of variance indicated a significant relationship between the superiority of small classes and the type of school attended. When, however, a more detailed comparison was made as between one size of class and another, analysis of variance suggested that there was not, overall, any significant difference between the average performances of pupils in classes of various sizes in these two tests.† This is associated with the figures obtained for classes of forty to forty-nine and those of classes of thirty to thirty-nine where the difference was significantly in favour of the larger classes.

A further inquiry into reading ability in a sample of fifty-one primary schools in Kent provided evidence as to the results of tests of reading-comprehension and (non-verbal) mental ability administered to rather more than seven thousand pupils of ages seven to eleven (Morris). Good reading attainment was found to be to some extent significantly associated with large schools, superior school buildings, urban areas, deliberate and early tuition in phonics and classes with an unfavourable pupil-teacher ratio (large classes).

Not dissimilar findings came from an investigation of much smaller scope in which information collected in brief visits to fifty schools there was some indication of little relationship between level of attainment, class-size and informal methods (Kemp). The last two variables were also studied in a year's experiment in the teaching of reading by a phonic-word method (Daniels and Diack). In this the experimental class of forty-four non-readers was contrasted with three control groups of twenty-five, twenty and ten pupils of similar reading ability and similar home background. At the end of the year significant superiority was shown by the larger class along with greater independence of attack on the part of its pupils.

Another type of indirect approach to the same problem of class-size is represented by a study of Lundberg in which a reduction in class-size from thirty-seven pupils to thirty pupils was (in a comparison with past records in the same school and in relation to national norms on the tests used) estimated to have led to an improvement in achievement, in attendance, in behaviour and in morale. This may usefully be contrasted with an extensive comparison of scores in standardised achievement tests of ability and personality administered in a city in Southern California (average size of class in Grade III: 29·5) and in British Columbia (average size: 38). In both cities scores were well above national norms but those from British Columbia compared favourably with those from

---

*Pidgeon, D. A. : Data on Class Size obtained from the Population Investigation Committee Survey (personal communication).

†Penfold, D. M. E. Personal Communication.

South California (Wilson). The significance of the differences was not, however, indicated in either of these studies.

Once again, therefore, it has to be said in the light of such incidental enquiries that class-size in itself does not appear to be a significant factor in achievement; and further support is lent to the earlier generalisations of Rice, Butsch, Otto and von Borgersrode that the secret of the difference between the success of one teacher and another must be looked for in something other than the mere size of the group under tuition.

Fuller Recognition of the Complexity of the Teaching Situation

Attempts to take account of other complexities in the teaching situation may be illustrated from Eastburn's extensive experiments with three hundred and sixty eleventh grade pupils arranged as one class of sixty and two classes of thirty at each of three ability levels in English and History (the size of the groups being reversed for the two subjects). These boys and girls of about the age of sixteen were initially matched by age and sex on a test of mental ability, on objective tests, and on their teacher's ratings. Each teacher took both large and small classes and followed the same type of teaching programme with each. At the end of a year, re-testing showed that the average scores of small classes on objective tests did not differ significantly from those of large classes in either English or History on the lower level and the upper level or in History on the middle level; but scores were significantly higher in the larger classes in English at the middle level of ability. A test of attitude given to all the classes showed no significant differences, and among the teachers who had taken part in the experiment there was no unanimity of opinion in their agreement with a list of the arguments commonly offered for and against large classes. Class size had no appreciable effect on disciplinary problems at any of the three ability levels; and the amount of inattention observed bore no relation to the size of the class.

In a second experimental study a year later, it was arranged that the same three teachers (with the aid of a secretary shared amongst them) should each take five classes of fifty pupils for English. The pupils were matched with pupils in the large and small classes of the previous experiment; and the findings again threw doubt on the assumption that pupils of lower ability should necessarily be taught in smaller groups than those of middle or upper ability. In the case of one teacher the learning of the five large classes was not significantly different from that of the groups of thirty in the previous year. With the second teacher the large classes did significantly better than the small classes at the middle level and equally well at the upper and lower level. In the case of the third teacher there was not a significant difference in achievement.

A somewhat similar enquiry by Cammarosano with young undergraduates in the field of the Social Sciences also contrasted groups of sixty with groups of thirty and showed comparable care in initial matching and in the covering of the same syllabus by the same professors over a full academic year. The results for tests for ' command of subject matter ' and for ' interest in public events ' showed no significant differences in mean scores in large and small groups; but in a test designed to estimate skill in applying principles the students from small classes in Sociology did significantly better than those from large ones while the size of the group did not significantly affect the response of the students taking Economics or Political Theory. In extensive interviews in which an attempt was made to discover the opinions of the students on such matters as vision, hearing, ventilation, opportunities for questions, discussion and consultation, the large groups appeared to have been as satisfactory as the small except for some complaints as to overcrowding. The teachers, on interview, expressed awareness of greater effort in establishing informality with large groups, in enlisting participation in discussion and in discovering special difficulties among students (because of the presence of an assistant taking the attendance) but they estimated that the time for marking papers, etc., was approximately twenty-five per cent. less for one class of sixty than for two classes of thirty.

In another study with nearly two thousand somewhat more senior University students, Husband compared the efficacy of lecture groups of one hundred and eighty to three hundred students as compared with smaller sections of thirty to sixty in the field of General Psychology. Again, the conclusion had to be that measurable differences were not statistically significant; and that the impersonality of the larger group did not lessen either the liveliness of questioning or the will to learn on the part of the larger groups; nor did it result in discernible gaps in their knowledge.

Research in Progress

Willingness to experiment not only with differences of size of groups but with differences in the use of non-teaching assistants is a feature of an important long-term study recently initiated in Jefferson County, Colorado. This, as reported by Patterson and by Romine, is interesting in that it proposes over a period of five years to study not only variations as between groups of ten, twenty, thirty-five, sixty and seventy High School pupils in different subjects, but also a variety of combinations of teams of teachers with and without clerical assistance. By this means it is hoped to test the hypotheses that in certain learning situations the number of learners is insignificant, that differing techniques are suitable for groups of different sizes, that when certain techniques are used the learning situation is unaffected by building facilities, that the morale of teachers working in teams with large groups is higher than that of those who work in isolation with small groups and that the problem of the shortage of qualified trained teachers can be met by the provision of non-teaching assistants.

Comparable concern with teaching techniques as well as with differences in measurable attainment is characteristic of a study in which in relation to the teaching of geometrical drawing in the first year of a secondary school two groups of seventeen (the small classes) and two groups of thirty-four pupils (the large classes) were compared. All four groups were matched for general mental ability and for age. The teacher, the classroom, the syllabus, the equipment and the time were the same for all; but one large group and one small group were seated according to their sociometric choices and in the other two groups no account was taken of the pupils' preferences and friendships. At the end of three terms, standardized tests showed no significant differences in attainment according to the size of the groups (using analysis of variance) but there were significant differences in the means in favour of the sociometrically-seated groups not only when each group was compared with another of the same size, but when the large sociometric group was considered against the small non-sociometric group.

III. Other Contrasting Fields of Educational Enquiry

In contrast to all this scarcity of evidence that class-size is in itself significant, mention may at this point be made of the findings which have accumulated in recent years as to the differences consequent on different types of climate in the teaching situation. Changes in this are known to have been effected without change in the number of pupils under instruction. They have been recorded at the primary level (Shimeld and Waite) and at the secondary level (Rowe) and in both controlled and uncontrolled situations.

Pupils respond differently to different teachers. The quality of their learning varies with the different approaches which can be used by the same teacher. It is modified also by changes in the type of school organization (streaming or non-streaming). It is altered by the encouragement of more spontaneous participation in group activities (competition or co-operation). It is transformed by changes in the self-picture of the pupils (by increases in their self-confidence and self-respect).

The differences to which attention requires to be paid from the point of view of those aware of the findings of educational research can therefore now with some assurance be said to include more than the externals supplied by administrative enterprise in the allocation of teaching load, size of class, time available and the like. The subtleties of inter-personal relationships have also a part to play. A school in which participation and co-operation is encouraged on the part of the staff is more likely to be one in which both teachers and pupils will develop high morale. A classroom in which a teacher permits opportunities for independent learning through co-operative activities (with provision of material for some measure of self-instruction and self-correction) is more likely to see the development of wholesome self-pictures and increasing ability to learn on the part of its pupils.

This does not mean that class-size is a matter of no importance. It serves rather as a reminder of the complexity of the teaching situation and of the supreme significance of the quality and the skill of the human being in charge of that situation. The size of classes has, with general approval, been reduced by law over the last sixty years; and discussion now centres on groups in the thirties and the forties, rather than the nineties or the hundreds. The protest of teachers against 'large' classes has been observably most vocal in periods of expanding enrolment—the 1920's and early 1930's, the late 1940's and again in recent months when the 'bulge' in birth rate is beginning to affect the secondary schools. This protest is not to be disregarded. It merits attention in terms of the effect of a larger group upon the teacher's load—of overcrowding, of lack of equipment, and of inadequacy of secretarial help in the cyclostyling of material and in the keeping of records. In interpretation of the general trend of theoretical discussion, public opinion, and research evidence over recent decades, it can, however, be said that there is no justification for supposing that a small class under the charge of an imperfectly educated adult is preferable to a larger class with a skilled teacher in command. In times of shortage of teachers or inadequacy of buildings, the mere addition of more adult ' bodies ' to function as ' teachers ' with smaller groups appears, therefore, to be less wise than attention to improved morale among skilled teachers through other use of auxiliary manpower along with more generous provision of teaching aids designed for the encouragement of active individual co-operation of pupils busily employed in the furthering of their own education.

Rice was feeling his way towards this when, in discussing scientific management in education, he dared to report the apparent irrelevance of size of class, duration of lesson and method of teaching employed. Scientific management in schools is now known to involve as a priority the establishment of satisfactory relationships—of a situation in which pupils and teachers are aware that in some sense they are appreciated, that some measure of success is available to them and that they have some understanding of the meaning of the activities in which they are asked to co-operate. Good buildings and space for working—the size of the most immediate group—are indirectly contributory to this as indicative of respect and appreciation from the more remote and impersonal ' organizers ' of institutional life. As only one of the variables in the teaching situation they cannot of themselves ensure that conditions within the classroom will necessarily be contributory to learning. In this there probably lies one of the reasons why, in the light of educational research, it has not proved possible to confirm the hypothesis that size in itself is a definitive issue.

Bibliography

ADOLPHSON, L. J. and UMSTATTD, J. G., " Teaching Load and Assignments." *Rev. Educ. Res.*, XIX, 3, 235-239. 1949.

ANDERSON, K. E., " The Relationship between Teacher Load and Student Achievement." *Sch. Sc. and Math.*, 1, 6, 468-470. 1950.

BAGLEY, W. C., " The Relation of Class Size to Effective Teaching." *School and Society*, 62, 1603, 164, 165. 1945.

BRADFIELD, L. E., and EDWARDS, T. W., " Employment Practices and Working Conditions in the Elementary and Secondary Schools." *Rev. Educ. Res.*, XXVIII, 3, 222-233. 1958.

BUTSCH, R. L. C., " The Teaching Load." *Rev. Educ. Res.*, IV, 3, 297-301. 1934.

CAMMAROSANO, J. R., and SANTOPOLO, F. A., " Teaching Efficiency and Class Size." *Sch. and Soc.*, 86, 2137, 338-341. 1958.

DANIELS, J. C., and DIACK, H., *Progress in Reading*. University of Nottingham Institute of Education. 1956.

DAWE, H. C., " The Influence of Size of Kindergarten Group upon Performance." *Child Dev.*, 5, 295-303. 1934.

DOUGLASS, H. R., *Organization and Administration of Secondary Schools*. Boston: Ginn. 1932 and 1945.

DOUGLASS, H. R., " Teaching Loads and Assignments." *Rev. Educ. Res.*, XIII, 3, 271-275. 1943.

DOUGLASS, H. R., and PARKHURST, A. J., " Size of Class and Teaching Load." *Rev. Educ. Res.*, X, 3, 216-221. 1940.

DOUGLASS, H. R., and ROMINE, S., " Teaching Loads and Assignments." *Rev. Educ. Res.*, XVI, 3, 255-261. 1946.

EASTBURN, L. A., " Report of Class Size Investigations in the Phoenix Union High School. 1933-34 to 1935-36." *J. Educ. Res.*, XXXI, 2, 107-117. 1937.

EASTBURN, L. A., " The Relative Efficiency of Instruction in Large and Small Classes on Three Ability Levels." *J. Exper. Educ.*, V, 1, 17-22. 1936.

GOODLAD, J. I., " Room to Live and Learn: Class Size and Room Space as Factors in the Learning-Teaching Process." *Childhood Educ.*, 30, 8, 355-361. 1954.

HARLAN, S., " Size of Class and Types of Instruction." *J. Educ. Res.*, XIV, 120-125. 1926.

HORN, E., *Methods of Instruction in the Social Studies*, pp. 68-78. Scribner & Sons. 1937.

HUSBAND, R. W., " A Statistical Comparison of the Efficacy of large lecture versus small recitation sections upon Achievement in General Psychology." *The Amer. Psychol.*, 4, 216. 1949.

IRWIN, M. E., " Size of Class and Teaching Load." *Rev. Educ. Res.*, VII, 3, 276-283. 1937.

JENSEN, M. B., and JENSEN, D. W., " The Influence of Class Size upon Pupil Accomplishment in High School Algebra." *J. Educ. Res.*, XXI, 2, 120-137. 1930.

KEATINGE, M. W., *The Great Didactic of John Amos Comenius*. London: Adam and Charles Black, 164, 165. 1896.

KEMP, L. C. D., " An investigation into environmental and other characteristics determining the attainment of children in a group of London Primary Schools." Ph.D. thesis, University of London. 1954.

LOCKE, J., *Some Thoughts Concerning Education*. London: National Society's Depository (n.d.).

LUNDBERG, L. D., " Effects of Smaller Classes." *Nations' Schools*, 39, 20-22. 1947.

Marcuse, F. L., " On Methods of Teaching Elementary Psychology." *J. Educ. Psychol.*, 42, 236-240. 1951.

Maxwell, J., *Educational and Other Aspects of the* 1947 *Scottish Mental Survey.* Scottish C. Res. Educ., XLI. University of London Press. 1958.

McDonald, D. and Umstattd, J. G., " Teaching Load and Assignments in Elementary and Secondary Schools." *Rev. Educ. Res.*, XXII, 3, 201-205. 1952.

McKeachie, W. J., " Students, Groups and Teaching Methods." *Amer. Psychologist*, 13, 10, 580-4. 1958.

McKenna, B. H., " Class Size and Teacher Load." *Nat. Educ. Ass. J.*, 46, 7, 437-438. 1957.

Miller, P. S., " A Quantitative Investigation of the Efficiency of Instruction in High School Physics." *J. Educ. Res.*, XIX, 2, 119-127. 1929.

Morris, J. M., *Reading in the Primary School.* National Foundation for Educational Research in England and Wales. London: Newnes. 1959.

Mowat, A. S., *City and Rural Schools: A Comparison of the Qualifying Candidates.* Scottish C. Res. Educ., XI. Univ. of London Press. 1938.

Olson, W. C., *Psychological Foundations of the Curriculum.* Unesco: Educ. Studies and Documents, 26, 61, 62. 1957.

Otto, H. J. and von Borgersrode, F., " Class Size," in Monroe, W. S., *Encyclopedia Educ. Res.*, 212-216. New York: Macmillan & Co. 1941 and 1950.

Patterson, G. E., *et al.*, " Classes of 10, 20, 35 and 70 under Varied Conditions are Taught in Jefferson County, Colorado, to Discover Effects on Students and Teachers." *Bull. Nat. Ass. Sec. Sch. Principals*, 42, 234, 165-7. 1958.

Peterson, E. T., *et al.*, " Teaching Load." *Rev. Educ. Res.*, I, 2, 92-98. 1931.

Rees, I. Gwynne. Presidential Address, Margate, 1949, N.U.T. Annual Conference. *Schoolmaster*, CLV, 2078, 477-479. April 21, 1949.

Rice, J. M., *Scientific Management in Education.* London: Harrap. 1915.

Rogers, D. C., " Pupil-Teacher Ratio in Chicago." *Nation's Schools*, 39, 43. 1947.

Romine, S., " Teaching Assignments and Instructional Loads in Secondary Schools." *Bull. Nat. Ass. Sec. Sch. Principals*, 42, 241, 55-63. 1958.

Romine, C., " Class Size." *Bull. Nat. Ass. Sec. Sch. Principals*, 42, 237, 278-9. 1958.

Ross, D. H., " Employment Practices and Working Conditions in the Elementary and Secondary Schools." *Rev. Educ. Res.*, XXV, 3, 227-238. 1955.

Ross, M. J., " Class Size in the Elementary Schools of Connecticut." *Res. Bull.* 1955-56.

Seashore, C. E., " Academic Business." *School and Society*, 62, 1603, 161-164. 1945.

Spitzer, H. F., " Class Size and Pupil Achievement in Elementary Schools." *Elem. Sch. J.*, LV, 82-86. 1954.

Tompkins, E., " Class Size: The Larger High School." Circular 305. Washington: Office of Educ. 1949.

Tompkins, E., " What Teachers Say About Class Size." Circular 311. Washington: Office of Educ. 1949.

Tompkins, E., " Enigma of Class Size." Washington: Office of Educ. (n.d.).

Thorne, E. H., " Watch Elementary Class Size." *Nation's Schools*, 37, 52, 53. 1946.

Watson, G., " An Evaluation of Small Group Work in a Large Class." *J. Educ. Psychol.*, 44, 7, 385-408. 1953.

Wilkinson, F. R., " Class Size in Higher Education." *J. Higher Educ.*, XXIX, 3, 149-157. 1958.

Wilson, J. A. R., " Differences in Achievement Attributable to Different Educational Environments." *J. Educ. Res.*, LII, 3, 83-93. 1958.

Wolfe, D., " The First Course in Psychology." *Psychol. Bull.*, 39, 9, 685-712. 1942.

For accessible material on changes in other aspects of the teaching situation, see:

Arbuckle, D. S., *Teacher Counseling.* Cambridge: Addison-Wesley Press Inc. 1950.

Corey, S. M., *Action Research to Improve School Practices.* New York: Teachers College. 1953.

Fleming, C. M., *Teaching: A Psychological Analysis.* London: Methuen. 1958. New York: John Wiley. 1958.

Redl, F. and Wattenberg, W. W., *Mental Hygiene in Teaching.* New York: Harcourt Brace & Company. 1951.

Rowe, A. W., *The Education of the Average Child.* London: Harrap (forthcoming). 1959.

Shimeld, A. and Waite, T., *Marking the Beacon Arithmetic.* London: Ginn (n.d.).

Wiles, K., *Teaching for Better Schools.* New York: Prentice-Hall Inc. 1952.

URBAN & RURAL DIFFERENCES IN ABILITY & ATTAINMENT

by F. Barr

Introduction

It is alleged sometimes, that children from the rural areas are on the average both educationally and intellectually inferior to their urban contemporaries. Often it is argued that this may be due to a general all-round inferiority of the rural population in educable capacity brought about by such causes as an unstimulating environment or a migration of the abler families to the towns. Alternatively, it has been stated that the rural schools, for reasons of size, equipment, and the training or experience of their staff cannot do as well for their pupils as the larger schools in the towns.

Such hypotheses have led to a number of suggested lines of action. For example, it has been argued that a higher or at least an equal proportion of grammar school places should be granted to children from rural schools as are given to urban pupils, irrespective of 'ability' as measured by tests. It has also been suggested that the teaching methods of the rural school should be changed to take account of a different kind of human ability and that rural and urban groups need separate treatment. Indeed, speculation has been more rife in proportion to the conflicting nature of the evidence than to its absence.

It is evidently important that administrators who have to decide on the amount and kind of educational provision which should be made in town and country and teachers who have to deal with children from rural or urban areas should have a clear idea of the facts. Even more important, in many ways, if such differences in ability and educational attainment do exist is it to know the nature of other differences which may be associated with them and their likely causes. Educational policy and, what is more important still, the teachers' attitude and method in the classroom can then be based upon something more concrete than vaguely supported belief.

Differences between the Means of Groups

(a) *Verbal Ability*

Much of the evidence comes from the results of surveys of different kinds, many of them with group tests of ability. Some of the earlier ones, such as the Scottish Mental Survey of 1932, indicated that there was little statistically significant difference between the mean scores of urban and rural children on the group verbal test used. Children were divided into three main groups corresponding to the geographical regions of Scotland. The investigators found that it was difficult to generalise about difference between the scores of rural and urban children. (Table 1.)

Table I

1932 SCOTTISH MENTAL SURVEY

MEAN DIFFERENCE OF I.Q. ACCORDING TO GEOGRAPHICAL REGION

Area	I.Q.
1. The Four Cities	100.86
2. Industrial Belt	99.86
3. Rural Area	100.19

It was difficult, too, to separate areas of town and countryside because of their inter-penetration, and rarely was there a true indentity of conditions present. Racial, social and economic factors combine in a variety of ways and were considered by the Survey to have some effect on the scores, but it was difficult to match different geographical areas which had similar environmental conditions. This study concluded by noting that as far as Scotland was concerned there was no significant inferiority in the scores of rural as compared to urban children; indeed, in certain of the more remote areas of the Highlands and Islands, there might well be isolated groups of people with higher than average ability.

Working in Northumberland, Thomson found similar isolated groups of gifted people in some of the more remote Cheviot valleys. Lewis, who studied the incidence of mental deficiency, also adds his warning that only an impetuous person would conclude, from the contemporary data that rural inhabitants as a group generally are inferior in mental endowments to the inhabitants of urban areas. However, Terman and Merrill, working in America, disagreed, and stated that rural children are definitely inferior in general ability as a group to urban children and claim that this conclusion is reached by other investigators. It was supported by Burt. Testing in rural and urban schools in England he found twice as many sub-normal children in the former as in the latter.

Post-war studies have tended to support the conclusions of Terman, Merrill and Burt rather than those of the Scottish survey. In 1947 a further survey was undertaken in Scotland and a ' follow-up ' of this was made in 1954. This time a small but statistically significant difference was found in the mean scores in verbal (' intelligence ') test scores of rural and urban children favouring the latter, a difference, however, less than the average difference between boys and girls throughout Scotland.

The 1954 ' follow-up ' survey of Scotland studied the question in detail and examined the extremes of high and low scores. The investigators found that whereas 46.5 per cent. of high scorers come from the cities only 29.7 per cent. came from the rural areas. Conversely, 55.7 per cent. of the low scorers came from rural areas as compared with 21.7 per cent. from the cities.

A similar conclusion was also reached by Cross and Revell in Cambridgeshire in 1955. They compared the performances of rural and urban children in a Local Education Authority's secondary school selection examination and found that whereas the percentage of rural children in the lower scores exceeded that of urban children this majority became progressively less until it was replaced by an urban majority in verbal scores (I.Q.s) of over 120.

Table II

CAMBRIDGESHIRE SURVEY, 11 + SELECTION TEST, 1955

PERCENTAGE DISTRIBUTION OF CANDIDATES GAINING HIGH VERBAL SCORES

Area	Verbal Test Scores				
	109 and under	110+	120+	130+	(N)
1. Rural	50.4	30.2	13.3	6.1	(655)
2. Urban	41.2	27.2	19.2	12.3	(739)
Difference (Rural compared with Urban)	+9.2	+3.0	−5.9	−6.2	

(*b*) *Attainments*

In 1952 the Bristol Institute of Education investigated the problems concerning the relative successes and failures of rural and urban candidates in the 11-plus selection examination of Wiltshire and found differences in Arithmetic, Verbal and English (Moray House Intelligence) tests. A comparison of the means showed that on all tests rural candidates achieved lower scores than their town contemporaries.

Table III

BRISTOL INSTITUTE OF EDUCATION: SURVEY OF WILTSHIRE URBAN AND RURAL PERFORMANCE IN 11+ SELECTION EXAMINATIONS

AREA	MEAN STANDARD SCORES			(N)
	I.Q.	English	Arithmetic	
Urban	100.60	100.23	100.68	1,126
Rural	97.69	98.58	98.55	1,304
Difference (Rural compared with Urban)	2.91	1.65	2.13	

(S.D. are all approximately 15, and a difference of 1.2 points is significant at 5 per cent.).
Source : Bristol Institute of Education publication No. 3, page 17.

This conclusion was also reached by the Cambridgeshire Survey in 1957, which found that mean standardised scores of rural children in English, Arithmetic and Verbal tests were significantly lower than those of the urban groups.

A similar result was found in a more intensive survey of sixty schools in the county of Kent, carried out in 1954-1958.. It showed that the scores of urban children in both reading and non-verbal tests were higher than those of rural children.

Perhaps the most convincing evidence, since it is based upon careful sampling of all children in all types of schools and areas in England and Wales comes from the National Foundation for Educational Research's own National Survey. The all-round superiority of urban children in the two basic subjects, and indeed in results on a non-verbal test of ability, is confirmed. How this shows itself in the most able groups is demonstrated by the following table which gives the proportions of candidates from urban and rural groups who scored 115 and above (population mean 100, and S.D. 15) on each of the three measures used.

Table IV

NATIONAL FOUNDATION FOR EDUCATIONAL RESEARCH: NATIONAL SURVEY. PERCENTAGES, URBAN AND RURAL, OF CANDIDATES WHO SCORED 115 AND OVER

TEST	URBAN	RURAL
	%	%
1. Arithmetic (problem)	17.7	10.0
2. Reading	15.8	9.7
3. Non-Verbal	17.9	13.4
Average ..	17.1	11.0

It appears that the difference in attainments between the rural and urban children is most marked in reading, a fact confirmed by the Ministry of Education surveys of reading reported in 1950 and 1957.

Thus the consensus of recent evidence derived from a variety of surveys undertaken by different workers and using different tests in various parts of the country is that the measured ability and attainments of rural children are lower than those of their urban contemporaries, and that this discrepancy is somewhat more marked in verbal tests and in tests of reading than in non-verbal tests and tests of arithmetic.

Incidence of Gifted, Average and Backward Children

It is important for teachers to know what proportions of gifted, 'average' and educationally backward children there are likely to be in their classes. Similarly, administrators will want to know the proportions of these groups which are to be expected in their areas. Both, then, can provide suitable education accordingly.

In the preceding paragraphs, and especially in Tables II and IV, some indications are given of the differential incidence of high scorers in the two broadly contrasted rural and urban groups. Other surveys provide similar approximate indications. Burt, for example, found that in rural areas there were nearly twice as many educationally subnormal children as in towns. The 1954 Scottish Survey also discovered differences; for every 5.7 low scorers on verbal tests in areas other than towns and cities, there were 2.7 in the cities whilst for high scorers the ratio was reversed: 3 : 4.7. In Cambridgeshire, over a period of ten years, 24 per cent. of the urban candidates in the 11-plus selection examination obtained Grammar School places but only 20 per cent. of the rural candidates gained similar awards.

It seems from these results, therefore, that teachers and educational administrators may expect important differences in the numbers of good, 'average' and below 'average' children. Indeed, we may hazard the statement that the differences in the proportion of high and low scorers may be almost as great as 2 : 1 in favour of the towns.

Possible Causes of the Differences

(*a*) *The Test Problems*

Most of those who have investigated the problem of urban and rural differences have warned against the simple ascription of the apparent inferiority of rural children to something inherent in the rural environment or the more 'stimulating' culture of the town.

In a search for causes we must first call into question the nature of the instruments of measurement used. It has, for example, been suggested that the tests themselves may be biased in favour of urban children who respond better or more quickly to them.

One Local Education Authority was concerned about this problem and asked the National Foundation to study it. The basic hypothesis of the enquiry was that if urban children were indeed more sophisticated in answering objective tests then one would expect the gain in score between two separate administrations of similar tests to be greater for rural than for urban children. This would result from the proportionately greater increase in sophistication reaped by the 'unsaturated' rural children over the 'saturated' urban children.

Samples of urban and rural children were given the same verbal ability test. This was followed at a later date by a similar test. It is known that coaching improves scores initially but that further coaching or experience with this kind of test does not result in an improvement beyond a certain limit. If the town children had had experience of this type of testing they would be nearer this limit than the country children who had had little experience in answering such tests. Thus, whatever the *total* scores, achieved by either group, it would be reasonable to expect that the *improvement* between the first and the second testing would be greater in the case of the rural than of the urban children. The results of this experiment, in fact, showed a slightly opposite trend.

Table V

NATIONAL FOUNDATION FOR EDUCATIONAL RESEARCH, 1958
URBAN AND RURAL DIFFERENCES IN MEAN RAW SCORES
BETWEEN TWO SIMILAR TESTS, VERBAL 8A AND 8B

AREA	MEAN RAW SCORE		MEAN GAIN BETWEEN 8A AND 8B
	8A	8B	
Urban ..	36·83	40·70	+3·87
Rural ..	30·97	34·22	+3·25

The differences are not large in terms of standardised score and do not reach significance. The differences in favour of urban children in mean performances between the first and second testing, however, are statistically reliable. Thus, the study legitimately concludes that there is no evidence to show that the country children are at a disadvantage when compared with the town children in answering this type of test.

The investigators conducting the Cambridgeshire Survey also studied this problem of bias from a different aspect. Each item used in the English and Arithmetic tests was examined separately. They concluded that there was no evidence to support the idea that the form of the examination paper was responsible for any differences.

We are obliged to conclude that neither the content nor the type of question used in the surveys previously mentioned favours one or other group. The measured differences, therefore, indicate a reality rather than an artefact of the measurement situation itself.

(*b*) *Migration*

If this is accepted then other explanations must be sought. The one most frequently advanced is that the abler families have moved and continue to move from the countryside to the towns.

Throughout recent history there has been a drift from rural to urban areas. In 1801 about 60 per cent. of the people of England and Wales lived in rural districts or in towns with less than 2,000 inhabitants; by 1931 this percentage had fallen to 20. The drift to the towns has continued even during the last generation, particularly from the more remote parts of the kingdom. This could explain why although Thomson in 1922 and the Scottish Survey of 1932 found groups of more gifted people in the isolated parts of the Cheviots and Highlands of Scotland, such pockets no longer appear to exist according to later work. These remnants of a vigorous peasantry may have largely migrated by now to the towns.

The 1947 Scottish Survey paid special attention to the problem of migration. It found that between 1936 and 1947, on the average, more children of under 11 years of age moved to ' Other Areas ' than vice versa. This apparently contradictory finding is explained by the unusual nature of the decade concerned, which covered the years of war when junior school age groups, in particular, were evacuated from the towns to the country. The research workers found, incidentally, that in the age groups over 11 years, the balance of the movement was the more usual one, from ' Other Areas ' to the towns and cities.

It seems probable that in general the families who leave the countryside are the more vigorous and adventurous ones. This seems to be confirmed by the finding of the 1947 Scottish Survey who compared the test scores of inward and outward migrants. The average score of city-born children migrating to ' Other Areas ' was 36.1 points; that of those who went to the towns from the countryside was 43.2 points. This, as the Report of the Survey points out, is a considerable difference. However, if there were a continual drain of talent from the countryside to the towns we should expect the gap between the abilities and educational attainments of town and country children to become steadily larger. This does not seem to be supported by the facts.

The apparent paradox is partly explained perhaps by the fact that mobility is not perfect; but there is an even more important factor. Although parents of high ability tend to have able children this does not occur invariably; some children are less gifted than their parents. Conversely, parents of poor ability do have children more able than themselves. Whatever the influence of the migration of the more enterprising to the towns, the countryside continues to produce talented children, some of whom may stay, although others subsequently migrate to the towns.

Migration thus affords at least part of a valid answer to the problem of why some early surveys found not only virtually little difference between rural and urban scores but also why some pockets of gifted people remained in remote districts. It explains why recent studies have found statistically significant differences. The above modifications of the crude hypothesis of migration can explain why the educational gap is not greater between urban and rural children. But it is unlikely to be the sole true answer.

(*c*) *The Problem of Country Schools*

There is a popular belief that small classes are better than large ones because they allow the teacher to give more individual attention to each child. Despite this, many parents of rural professional groups prefer to send their children to town schools often at considerable expense.

The Cambridgeshire Survey studied this paradox by constructing a table comparing the number of grammar school places obtained by rural pupils according to the size of school which they attended. The results suggest that the smaller schools, mainly rural, obtained fewer places than might have been expected, whilst the larger schools, which were situated mainly in the towns, obtained more places.

Confirmatory evidence is derived from two surveys undertaken by the National Foundation. At first, data compiled by the Population Investigation Committee appeared to support the view that smaller classes in the rural schools were favourable because they allowed more individual attention, but further analysis revealed that the smaller classes occurred mainly in private schools which were largely patronised by parents of professional occupations. This factor, rather than the size of classes, appears to be the important one.

The Kent Reading Survey found that not only were the scores of rural children poorer than those of their urban contemporaries on non-verbal and reading tests but that the high scores were typical of large schools which had good buildings and junior but no infant departments. This type of school was found to be typical of the towns rather than of the countryside, and particularly of the new housing centres of good socio-economic character. Evidence also was found which seems to suggest that larger classes are slightly more favourable to formal attainment in reading than the smaller ones, but this is a tentative conclusion on which much more evidence is needed, and it should not be interpreted to mean that large classes lead to better attainments invariably.

(*d*) *Socio-Economic Factors*

The 1947 Scottish Survey drew attention to the fact that when children were grouped according to the occupational category of their parents, irrespective of whether they came from rural or urban environments, differences related to socio-economic groups appeared. Of the children from parents of professional occupations, 9 per cent. fell in the high scoring group as compared with 4 per cent. of those from agricultural worker's families. Both this and the 1954 'follow-up' suggested that the superiority of town children was likely to be due less to any marked differences in the cultural opportunities of the urban or rural environments as such than to immediate family living conditions.

It will be recalled, too, that differences are most marked in verbal abilities. Describing a locality in the Pennines in which groups of miners, mill-workers and farmers were juxtaposed, Semens noticed differences in vocabulary whereby a child of an urban mill-worker would find difficulty in underlining the misfit in the following series of farming terms:

'stirk, gilt, reckling, hoggett, colt.'

These words, of course, come from a highly technical vocabulary and a similar series could be produced from almost any craft or trade, urban or rural. What is important, however, is to note the fact that differing family environments are likely to have not only differing vocabularies but vocabularies of a differing degree of richness.

This is in itself the superficial sign of more fundamental basic differences. Vocabulary represents a store of concepts; it is related to ways of life and to the richness or poverty in verbal-intellectual stimulus which the immediate environment provides for children in their earlier years. Since much of what we understand as learning ability as shown in primary and secondary school depends upon verbal-intellectual development, we should expect no small relationship between the educational and occupational level of the parents and the measured abilities of their children.

An Australian Survey of Socio-Economic Influences

Evidence exists that these conditions do occur in Australia. Urban and rural conditions were studied by Oeser, Emery and Hammond in 1949. The authors indicate that the educational environment in both rural and urban schools was closely similar since the pupils are similar ethnically, the compulsory school age fixed, the administration centralised and hierarchical and the text-books common to all State schools. Teacher training is centralised and uniform in the various colleges, and in general standards are supervised by inspectors attached to the State Education Departments. In spite of this, however, the authors found that rural children were poorer both in ability and attainments.

The Survey suggests that the differences are partly due to differences in ideology. Teachers are trained generally in an urban environment and a candidate from the country must abandon his rural background to a large extent thereafter adopting the outlook and social position of the urban professional groups.

Nor were teachers the only ones to be affected by differences of ideology because parental attitudes to school appeared to be dissimilar. In the rural areas it was found that 87 per cent. of parents disapproved of homework whilst in the smaller towns the figure was 67 per cent. In the city only 24 per cent. objected although as the authors pointed out this figure was probably biased because the parents knew that the interviewers were associated with their children's school and thus may have given what they considered to be ' acceptable ' answers.

Homework, too, was found to be a major source of tension between school and home. Parents disliked being measured by the standards of the teacher. This dislike would be less where, as in the towns, more of the parents were of similar social backgrounds and educational level to their children's teacher. Parents of an agricultural background might be tempted to disparage ' book-learning ' openly, although they may secretly respect or envy it.

Whereas urban middle-class parents tended to exert pressure on their children to do well at school because they regarded education as a means to achieve status and upward social mobility, farm parents tended to perceive higher education as a ' frill '. They believed this ' frill ' to be a good thing for their daughters as a social rather than as a vocational asset but considered it to be inessential for their sons.

Attitudes towards the curriculum differed. One-half of the town parents preferred the ' Mathematics-English-Social Studies ' subjects, but only a quarter of the rural parents agreed and instead a half of them preferred the ' handwork ' subjects. Each group, in fact, probably favoured a particular side of the curriculum for its apparent vocational value in their environments. Their children were aware of their parent's preferences and this awareness may well have influenced their attitudes towards, and hence performance in, the different school subjects.

Socio-Economic Influences in Britain

Australian and British conditions are similar. The 1947 Scottish Survey noted that high scores correlated highly with socio-economic factors such as a low average number of people occupying each room and with small families; conversely, low scores were found to be frequent among large families and with a high occupancy rate. A high score and its correlates proved to be typical of professional groups whilst unskilled and farm-labour groups were associated with a low score and its correlates.

Confirmatory evidence is given by Floud, who investigated the patterns of grammar school success in the County Borough of Middlesbrough and the South-West Division of Hertfordshire. The former is an area of heavy industry in whose economic history boom has been followed by slump, whilst the latter area has enjoyed longer periods of prosperity. The inquiry found a close relationship between ability, as measured by verbal tests, and opportunity, as measured by entrance to grammar schools, at all social levels. When those children of clerical workers who had been given grammar school places, were combined with the similarly successful children of manual workers there was found to be little difference between the two areas, but large differences were noticed when the ratios of the two socio-economic groups were compared. Thus, in Middlesbrough the ratio of grammar to non-grammar school entrants in the clerical groups was 1 : 3 and in the manual groups 1 : 8, whilst the respective ratios in S.W. Hertfordshire were 1 : 2 and 1 : 7.

The inquiry also discovered the interesting fact that it was not so much the actual socio-economic status as measured by income which was the critical factor. The most successful children certainly came from families which were smaller and better housed, but also success was found to be more frequent where there were better educated parents who were favourably disposed towards and interested in the education of their children.

The Influence of the Home Environment

At first sight one finding of the Cambridgeshire Survey appears to be contradictory. The investigators found that once rural children entered the grammar schools their performance was more consistent from year to year and easier to predict than that of urban entrants. The Report speculated as to whether urban candidates had their attainments pushed precociously in the primary school beyond their true capacity for grammar school work, thus leading to an unstable subsequent performance. Coaching for tests within schools can be forbidden or controlled by the authorities, whilst out-of-school coaching is probably a reflection of parental interest and anxiety which is perhaps stronger in the professional and urban artisan groups than among agricultural workers who form the majority of the rural population. The child of the manual worker in town or country may belong to a family whose attitude towards post-primary education is antagonistic or at best passive. This may be one reason for some of the unstable and disappointing grammar school careers.

Regularity of attendance at school and the length of a child's school life are two reasonably objective indices of parental interest. The 1947 Scottish Survey accepted 90 per cent. as a satisfactory average attendance at school. The investigators found that only 50 per cent. of rural children reached this figure as compared with 80 per cent. of children of small employers who are probably living in urban districts.

That children from professional families are encouraged to remain at school longer than those from unskilled workers' homes is supported by the figures in Table VI cited from the Central Advisory Committee's Report on 'Early Leaving'.

Table VI

CENTRAL ADVISORY COMMITTEE, REPORT ON EARLY LEAVING
LENGTH OF SCHOOL LIFE, GRAMMAR SCHOOL BOYS AND GIRLS

LAST FORM AT SCHOOL	PROFESSIONAL AND MANAGERIAL GROUPS		UNSKILLED GROUPS	
	Boys	Girls	Boys	Girls
	%	%	%	%
1. 6th	47	41	7	6
2. 5th	46	53	55	51
3. Below (at 15+)	7	6	38	43

(*Source* : Table P.)

The Report considers that one of the most important influences on the length of a child's school life is the home background. It states that the results of overcrowding are particularly serious. This handicap apparently becomes aggravated in the later years of school life and tends to falsify predictions of success made in the earlier years of the school career. Shortage of money by itself was not found to be the most common reason for early leaving; the general conflict of different social ideologies between the child's school and home environments often results in tensions which influence both the child himself as well as the parent to decide to end the secondary school life prematurely.

(*e*) *The Population Investigation Committee Data*

One of the most extensive post-war studies which provides information about rural and urban differences is that conducted by the Population Investigation Committee. Originally this investigation began as a medical and social inquiry into the development of children born during one week of March, 1946 throughout Great Britain. From the time these children entered school their educational development has been studied by the N.F.E.R. which has administered Reading and Non-Verbal tests to a sample of over 5,000 children.

When the results are analysed, rural children are found to be lower scorers than the town children if the groups are separated according to where they lived. When, however, the analysis is carried further and pupils are grouped according to the occupational status of their family, the crude rural-urban differences tend to disappear. A comparison of the urban and rural professional groups reveals that the difference between their mean scores is negligible; a similar comparison of rural and urban manual groups likewise shows little significant difference. However, when the mean score of the combined urban and rural professional groups is compared with that of all those employed in manual occupations there appears a significant difference with the former achieving the higher mean scores.

Thus, in common with the other large-scale surveys already considered the P.I.C. data reveals a strong positive correlation between test scores and socio-economic environment, not only in parts of the country but throughout the country as a whole.

Conclusion

From all that has been stated, it is clear that as a group, children from rural schools are likely to have lower mean scores on verbal tests and on tests of the basic school subjects than their urban contemporaries. Practically, this means that there are likely to be proportionately fewer academically able and more academically weak children in the countryside than in the towns and suburbs. However, since the main influence at work is not the nature of the rural or urban environment itself but rather the structure of the socio-economic group from which any school draws its pupils, there will be exceptions to this general tendency: moreover, any reversal of the trend, for example, for the abler and more enterprising families to migrate to the towns, will bring about a quite different state of affairs. In terms of educational provision it is clearly necessary for an educational authority to have a rough estimate of the different mental make-up of urban and rural groups and at the same time to have some idea of past and present population trends. It seems likely, too, that in the past some rural schools have been too much under the influence of urban ideologies, educational method and content on the one hand and less well provided materially than city schools. If this has coincided with a larger proportion of dull pupils for whom no special provision is made, there has been something of a handicap for pupils and for teachers.

The most important single factor, however, associated with differences in ability as revealed by these studies is that of the level of encouragement a child receives at home—a factor related to the socio-economic level and education of his parents. Mothers and fathers who are really interested in their child's education and who maintain this interest throughout his school career seem to contribute greatly to his success. Even in relatively adverse material conditions this seems to apply; conversely apathetic or antagonistic parents constitute a grave handicap even for the ablest child.

This, however, raises two problems on which practical experiment and research are clearly necessary. We know little as yet about the intimate ways in which 'parental encouragement' affects the growth of children intellectually. Of more immediate concern, however, because it lies in the teacher's own province, is the practical question of what the school can do in some measure to compensate for any deficiencies in the child's home environment once they are manifest. Here there is room for the practicing teacher to carry out experiments which can only be of value.

Select Bibliography

CROSS, G. R., REVELL, K. J., "Note on Grammar School Selection in a Rural Area." *Bulletin No. 7*. National Foundation for Educational Research, March, 1956.

DIETZ, C., *La Selection des eleves de l'enseignement secondaire* (*Recherches de statistique mathematique et de psychologue experimentale*). Paris, 1934.

FLOUD, J., "Social Class and the Eleven Plus." *Journal of Education*, Vol. 88, October, 1956.

INSTITUTE OF EDUCATION (University of Bristol), *Studies in Selection Techniques for Admission to Grammar Schools*: Publication No. 3. U.L.P., 1952.

LAHY, J. M., "Psychology Applied to Educational and Social Problems in a Rural District." *Occupational Psychology*, Vol. XII, No. 2, 1938.

LEWIS, E. O., *Report of the Mental Deficiency Committee (Part IV)*. London: H.M.S.O., 1929.

McMEEKIN, A. M., *The Intelligence of a Representative Group of Scottish Children*. Scottish Council for Research in Education. U.L.P., 1937.

MACPHERSON, J. S., *Eleven-year-olds Grow Up*. Scottish Council for Research in Education. U.L.P., 1958.

MINISTRY OF EDUCATION. 1. *Early Leaving*: A Report of the Central Advisory Council for Education (England).

London: H.M.S.O., 1954.

2. *Reading Ability:* Pamphlet No. 18. London: H.M.S.O., 1950.

3. *Standards of Reading,* 1948-56*:* Pamphlet No. 32. London: H.M.S.O., 1957.

Oeser, O. A., Emery, F. E., *Social Structure in a Rural Community.* Routledge and Kegan Paul (for the Department of Psychology, University of Melbourne), 1954.

Oeser, O. A., Hammond, S., and others, *Social Structure and Personality in a City.* Routledge and Kegan Paul (for the Department of Psychology, University of Melbourne), 1954.

Scottish Council for Research in Education, *Social Implication of the* 1947 *Scottish Mental Survey.* Scottish Council for Research in Education. U.L.P., 1953.

Semens, H., " Rural Reorganisation." *Journal of Education,* Vol. 90, March, 1958.

Stern, W., *Die Intelligenz der Kinder und Jungendlichen.* Berlin: 1928.

Straker, D., " Intelligence Testing in a Rural Area." *Occupational Psychology,* Vol. XIII, No. 1, 1939.

Terman and Merrill, *Measuring Intelligence.* London: Harrap, 1937.

Thompson, G. H., " The Northumberland Mental Tests." *British Journal of Psychology,* Vol. XII, No. 3, 1922.

Teaching Children to Read—II

THE RELATIVE EFFECTIVENESS OF DIFFERENT METHODS OF TEACHING READING

B. *The Place and Value of Whole-word Methods*

by JOYCE M. MORRIS

I. Introduction

DIFFERENCES in terminology impede any attempt to evaluate scientific evidence on the relative efficiency of different methods of teaching reading. There is, however, some agreement among authorities that most reading methods may usefully be classified into two broad groups on the basis of the psychological processes presumed to be involved. 'Synthetic' is the term frequently used to describe those methods which build up a recognition of wholes (words, phrases, and sentences) from parts (letter-names and sounds), and 'analytic' refers to those which begin with a recognition of wholes as units of meaning from which their constituent elements are broken down. Thus, the alphabetic, syllabic and all phonic methods constitute a synthetic approach to reading whilst whole-word methods, which include the word, phrase, sentence and story methods, may be defined as analytic. Alternatively, the latter are sometimes described as 'global', 'syncretist' or 'ideo-visual'.

On the whole, educationists in English-speaking countries today consider that the synthetic work involved in the first group of methods is too difficult for beginners in reading and is in any case 'psychologically unsound', being contrary to the principles of child development. Consequently, they advocate an analytic approach after children have enjoyed an informal period in which to develop confidence, an interest in books, and the necessary desire to read without which any method is useless. Later on, when a sight vocabulary has been acquired, instruction in breaking down words into smaller units is recommended in order to give pupils a technique for deciphering new words.

Research into the place and value of phonics in the reading programme (see first article in this series) would appear to support the view that whole-word methods should precede any systematic phonic teaching, for, in most studies, a successful use of synthetic methods has been shown to depend upon the level of maturity of the pupils. Thus, intelligent children, or those from home backgrounds rich in linguistic resources, will generally have reached an adequate stage of phonic readiness in the reception class, whereas for those who are dull or handicapped in any other way, systematic phonic instruction may desirably and safely be postponed until later in the school course.

It should be pointed out in this context that the successful use of a synthetic approach to reading by some non-English-speaking countries, particularly those of Scandinavia, does not contradict this idea that analysis should precede synthesis in the teaching of reading, since children in these countries start school at least one, often two, or even three years later than their English contemporaries.

The value of whole-word methods cannot, however, justifiably be inferred solely from a study of the use of phonic methods: the case for analytic methods of teaching reading must be examined in its own right. In doing this one may look at the historical development of the notions concerned and their embodiment in current educational practice.

II. Historical Background

Comenius is usually credited with the introduction of the word method. His *Orbis Pictus—The World of Sensible Things Pictured*, published in 1657, was, according to Monroe, the first illustrated textbook for children. In this he argued that when words are presented along with pictures representing their meaning " reading cannot but be learned—and indeed too, without using any ordinary tedious spelling—that most troublesome torture of wits."

It was not until the beginning of the 19th century, however, that Comenius' revolt against the established synthetic methods of teaching reading received strong support from educators. In France, Jacotot, advocating the use of visual methods, too, suggested that words rather than letters should form the child's introduction to reading. He even went so far as to state that, if a child learns by heart a sentence or a page of something, he will divide it for himself into words, syllables and letters, and be able to recognise them later in a different context.

About the same time, Worcester, in America (quoted by Smith), was writing in his *Primer of English Usage:* " It is not, perhaps, very important that a child knows the letters before he begins to read. It may learn first to read words by seeing them, hearing them pronounced, and having their meanings illustrated, and afterwards it may learn to analyze them or name the letters of which they are composed." His compatriot, Horace Mann, shared this view of children's development and strongly criticised the current letter-by-letter system of teaching reading, preparing the way for the first (mid-century) American readers specifically based on whole-word methods.

Froebel and Dewey indirectly, and Vogel with his ' method of normal words ', made further contributions to the ' whole-before-part ' school of thought. It was Decroly in Belgium, however, who, at the close of the century, carried their ideas to a logical conclusion by insisting that as one reads primarily for meaning the sentence is the appropriate unit for beginning reading.

Decroly exerted a strong influence on the development of whole-word methods of teaching reading. Indeed, it is often asserted that, except for the fact that children now are encouraged to handle books right from the start, the teaching in infant classrooms is basically Decrolian in character. The climate of educated opinion at the end of the first world war, or even before it, favoured a less mechanical approach to reading than had hitherto been in vogue. The difficulties of a phonic approach for very young children had become apparent. Many pupils, for example, stuck at the point of being able to sound out the letters but were unable to synthesize the individual sounds into words. The letter-by-letter peering at words encouraged by this method was neither conducive to speed nor fluency. Moreover, when giving phonic instruction, teachers found it difficult to make themselves heard at the back of a class without adding a distinct vowel sound to some letters, so that ' p ' became ' per ', ' f '—' fer ', and so on.

The weaknesses of synthetic methods seeming thus to be exposed, it is not surprising that the ideas of educators like Decroly exercised an increasing influence on teaching practice. Unfortunately, as is often the case, the pendulum swung too far in the other direction—so much so that, according to Gray, American teachers during the late twenties and early thirties were advised to consider meaning as almost the only factor in word perception and to expect their pupils to identify new words by guessing from the context.

Experienced teachers are rarely persuaded to change their methods by the voices of articulate radicals alone and many continued to give systematic instruction in word-analysis skills. In the middle thirties, however, some of the ideas of the Gestalt psychologists on visual perception began to penetrate to the schools, and appeared even to some of the more sceptical teachers to give psychological colour to practice in the ' progressive ' schools.

The fundamental hypotheses of the Gestaltists will be discussed later: here we may remark that recently experimental psychologists have tended to discard much of this early work. Vernon, for example, who in 1937 devoted considerable space in her book *Visual Perception* to Gestalt psychology, stated seventeen years later:

> " The apparently clear outlines originally drawn by the Gestalt psychologists have become blurred and dimmed. And it has seemed perhaps that their experimental results and their theories did not after all indicate a fundamental basis for the understanding of the nature of perception, but rather a one-sided exaggeration of certain features by no means the most important in perceiving as we ordinarily experience it."

In contrast, educationists generally have continued to stress the support given to whole-word methods by the Gestaltists. The following statement from *Backwardness in Reading*, by Duncan, is typical of many:

> " We tend to see a whole word-pattern before we see the parts of it. A young child may recognize the word ' elephant ' long before he knows the component letters. It is a distinctive ' word-pattern '. Even the most backward older pupils, who do not know letter-sounds, although they may have been ' drilled ' in them for years, recognize the words ' ice-cream ', ' orange ', and ' lemonade '. This tendency to see in ' wholes ' is emphasized in the *gestalt* hypothesis."

One wonders how such conflicting opinions can exist side by side until it is realised that psychologists interested in an experimental approach to visual perception are concerned mainly with the fundamental Gestalt hypotheses themselves, whereas educationists usually concentrate attention on practical consequences which may derive from them. If teachers are to feel confident in the use of whole-word methods for beginning reading, a re-examination of the part played by Gestalt psychology in their development is very necessary.

Strengthened by the rise of the ' experience ' approach to learning in the thirties, Gestalt theories undoubtedly influenced development towards greater attention to word meanings and contextual clues. The look-and-say method of individual word recognition was ' old ' in 1933, according to a Board of Education report. By then the sentence method for beginning reading, which implies not merely a single, relatively uncomplex visual gestalt but a much more complex, contextual and meaningful one in which visual clues do not dominate, had taken its place in many English schools. Educational practice had, in fact, far outrun such evidence as was provided by the actual experimental data of the Gestalt psychologists.

As might have been expected, there was a certain amount of criticism of this development in teaching practice, but not until the war years, when the extent of illiteracy among the Armed Forces was revealed, did it become serious. Public opinion, aroused by the

Press, virtually demanded that educationists should re-assess the validity of the methods by which reading was taught in primary schools. A series of investigations by local education authorities and the Ministry of Education in England certainly produced evidence of a deterioration in children's reading standards. This was ascribed largely to the disruptive effects of the war but other contributory causes were suggested. There was no recommendation of a return to early systematic phonic instruction and no general attack on the use of analytic (either word-whole or sentence) methods for beginning reading apart from a criticism of their exclusive use. Inquiries in America and other countries resulted in similar conclusions.

If indeed teachers follow the advice of authorities on the teaching of reading, the word-whole method, or the later sentence approach of the thirties, should predominate in the reception classes of our schools. How far this is a fact must be judged from recent inquiries.

III. Whole-word Methods Today

A clear picture of teaching practice in any large area is difficult to obtain. Even in one school, long classroom observation is necessary before a refined and reliable assessment of the methods used in the early stages of teaching reading can be made. It is misleading, too, to speak of different reading methods as though each were self-contained, and as if teachers confined themselves to one particular method, and are consistent in their approach for all children. Nevertheless, broad trends towards an analytic or synthetic approach may be distinguished.

According to Duncan in an appendix to *Language*, a 1954 publication of the Ministry of Education, " in general, in the United States of America during the past twenty-five years the use of phonic work in teaching reading has disappeared. No widely circulated readers published in that country in recent years has any phonic basis nor contains any phonic work." Russell's inquiry, described in a previous article, would suggest that this is a rather exaggerated view of the situation. It is, however, probably true to say that analytic methods predominate in American schools today, at any rate in the early stages of learning to read.

In this country as a whole no recent survey of current trends has been carried out, and the opinions of educationists conflict to some extent. Duncan, for example, states that synthetic methods have been out of favour in English schools for some years, and Daniels and Diack appear to support his statement. Gagg, on the other hand, from a knowledge of over 400 schools, challenges their observations and points out that, whilst whole words usually form a child's introduction to reading, " infant teachers do in fact use ' phonics ' and that in almost all the popular published reading schemes phonic practice is found."

Conclusions about teaching practice in England, drawn either from individual observers or from a study of one county, are obviously of doubtful general validity. There is, however, some evidence from the Foundation's recent Kent inquiries (Morris) which backs up the description of present practice given by Gagg. In these studies, as will be seen in Tables 1 and 2, a distinction has been made between the ' word ' and ' sentence ' methods, although both represent stages in the evolution of analytic methods.

Table 1

DISTRIBUTION OF READING METHODS USED IN THE INFANT DEPARTMENTS OF SIXTY KENT SCHOOLS

Methods*	Urban Schools	Rural Schools	Total
A+B+C+D	0	4	4
B+C+D	32	15	47
B+C	2	7	9
TOTAL	34	26	60

*A—Alphabetic. B—Phonic. C—Word. D—Sentence.

All schools in this random sample used both analytic and synthetic methods of teaching reading, although a tendency for rural schools to favour the latter may be observed in the tables. In four village schools, experienced teachers, having obtained satisfactory results with a previous generation by the alphabetic method, continued to use it in conjunction with other methods. Teachers in nine schools, also of rural type in spite of the fact that two were in districts administered by an urban authority, considered that the sentence method was not particularly suitable for 'the slower type of child attending a country school'. They firmly believed that, if pupils rely too much on contextual clues, they do not gain skill in recognising individual words. Children in these schools were, therefore, taught mainly by the phonic and word methods, and sentences were only introduced at a third stage when some facility in word building and recognition had been acquired. Similarly, in the more important choice of methods for beginning reading (Table 2) a sizeable proportion of rural schools appeared to be disregarding modern educational theory by using a synthetic approach.

Table 2

DISTRIBUTION OF METHODS USED IN COMMENCING TO TEACH READING IN SIXTY KENT SCHOOLS

Methods	Urban Schools	Rural Schools	Total
Sentence	12	8	20
Word	10	8	18
Phonic	5	8	13
Sentence + word	3	2	5
Sentence + word + phonic	4	0	4
TOTAL . .	34	26	60

This survey in Kent schools clearly demonstrates the freedom given to English teachers to decide the manner and content of their teaching. At the same time it gives evidence of the continued popularity of word and sentence methods for beginning reading.

A similar bias in favour of analytic methods is also revealed in a report published in 1949 by the International Bureau of Education on the teaching of reading in forty-five countries. Apparently, wherever synthetic methods are still widely used, it is often the form of the language which dictates this practice. In Burma, for example, " the particular formation of the Burmese letters is one of the chief causes of difficulties in learning to read, and would appear to force teachers in many schools to use the spelling method." Regularity of spelling is not always conversely a guarantee that synthetic methods will be adopted for, although Turkish is one of the world's most regularly spelt languages, schools in Turkey have abandoned this practice in favour of analytic methods. Moreover, a non-synthetic approach is making headway in France and Germany, despite the fact that French and, particularly German, are more regularly spelt than English.

That analytic methods are widespread does not mean that they meet with unqualified approval in all quarters. Diack, for instance, states that there is little scientific justification for what he considers to be the most frequently quoted observations in favour of ' visual pattern—whole-word theory ', namely, that " fluent readers skim along and do not pause to analyse words into letters," and " children recognise some words by their general shape." Having examined and dismissed some of the experimental work on which this theory rests, he concludes that " ' look and say ' derives its strength from at least two other sources: (1) apparently quick results—but how deceptive! (2) the ease with which reading material can be produced when any word will do."

Replying to this attack on whole-word methods Whitehead points out that, in this country at any rate, no one has contended that children can complete the process of learning to read by these methods alone. Even ardent supporters of analytic methods, like Boyce, give detailed advice about training in word analysis, although such instruction is generally regarded as more appropriate for first-year juniors than for infants.

In Whitehead's opinion the arguments in favour of visual pattern—whole-word theory referred to by Diack have more cogency than Diack recognises " since (1) it is helpful to future progress to establish from the beginning habits of eye-movement along the line of print which are similar to those used by the mature reader, and (2) there is clear evidence that children develop the visual power of discriminating between whole words earlier than that of discriminating between individual letters." Furthermore, Whitehead considers that Diack has omitted to mention two considerations which are usually regarded as far more decisive. First, through analytic methods, particularly the sentence method, " the child from the beginning practices reading as a process of obtaining meaning from printed symbols—and that after all is the essence of reading." Secondly, " because of this (and because of the *comparatively* wide range of words which can be used under this method) it becomes much easier to maintain the purposeful interest and desire to learn which is all-important for young children."

Whitehead's comments and those of other educationists show that the dispute about analytic methods centres on one main question. Should whole-word methods be used to *introduce* children to reading? In attempting to answer this question from the results of research, it is necessary not to lose sight of the fact that the teaching of reading cannot be treated merely as a technical problem in linguistics, but must be considered in its broader, human context of the classroom environment appropriate to young children when they start learning to read.

IV Psychological Rationale of Whole-word Methods

(a) Early experiments

Our knowledge of the psychological processes involved in reading is based mainly upon laboratory experiments conducted during the latter half of the nineteenth century. Until this time it was generally believed that words are read by compounding the letters. In 1885, however, J. M. Cattell dealt a severe blow to the, then widely used, alphabetic method by demonstrating that words are not read by spelling them out.

Using a tachistoscopic technique he found (according to Woodworth and Schlosberg) that, with a very brief exposure under adequate illumination, his subjects could apprehend equally well three or four unconnected letters, two unrelated short words, or four connected short words. Since even the two short words, not to speak of the four-word sentence, exceeded the letter span, and since a familiar short word could be read as quickly as a single letter, Cattell reasoned that the reader's basis for word recognition must be the ' total word picture '.

Three years later Erdmann and Dodge confirmed Cattell's findings. They reported that, whereas the span for unconnected letters was only about four or five, familiar words, even as long as twelve or twenty letters, were correctly read from a single exposure of one-tenth second. From this they inferred that the ' general word shape ' must be the primary cue for word recognition with a few clear letters close to the fixation point as supplementary cues.

About the same time, studies of the behaviour of the eyes in reading invalidated an earlier assumption that they move continuously along the line. It was discovered that the eyes execute a series of short, quick movements, fixating at several different points (either words or groups of words) along each line according to the reader's familiarity with the material, its difficulty, and its interest for him.

These early researches supported the newer concept of reading for meaning put forward at that time by educators such as Decroly. They are frequently quoted in the textbooks for teachers as providing evidence of the advisability of teaching reading by whole-word methods. Yet, it must be remembered that much of this early work was carried out with adults who could already read. As Diack points out: " If evidence drawn from adult readers gives validity to the argument that children ought to be taught by ' visual pattern ' means, then surely it is also valid to say that because an adult pianist sees and plays whole chords and phrases automatically there is no need for beginners to do five-finger exercises or scales."

Advocates of whole-word methods for beginning reading might challenge their critics in a similar manner, for a good deal of the evidence in support of early systematic phonic instruction stems from work with older, backward children. It seems equally invalid to deduce that methods which are eventually effective with pupils who have already failed to learn to read as infants are appropriate for young children at the very beginning of their school career.

Likewise those who quote the original work of the Gestaltists in support of analytic methods may be criticised, since the early findings of this school of psychology were founded not on the visual perception of printed symbols, i.e. letters and words, but mainly on optical illusions. Furthermore, as in the pioneer experiments referred to above, except for a few studies of mathematical and quasi-mathematical problems, adults were used as subjects and not children, although deductions about the latter were frequently made.

There appears to be a strong case for limiting discussion on the psychological rationale of whole-word methods to evidence from research with children. However, as already stated, Gestalt psychology played such an important role in the development of analytic methods that a brief examination of some of its basic concepts is called for first of all.

(b) The influence of Gestalt psychology

In 1912, when phonic methods were still flourishing in the schools, three German psychologists, Wertheimer, Koffka and Köhler, openly expressed their discontent with the current associationist, or ' bricks-and-mortar ' psychology, and so laid the foundation of the school of Gestalt psychology. They considered that human experience has a natural unity or ' wholeness ' and is more meaningful than had hitherto been suggested. Thus, they argued, we must not work ' from below upward ', i.e. building a whole from its parts, but rather ' from above downward ', so that starting with the meaningful whole we increase our understanding by finer and finer discriminations of the parts. Ellis, translating from the German, defines a ' gestalt ' as " a whole the behaviour of which is not determined by that of its individual elements but where the part-processes are themselves determined by the intrinsic nature of the whole." Accordingly, the main aim of Gestalt psychology was to determine the nature of these wholes which are ' more than the sum of their parts '.

Unfortunately, in pursuing this aim the reading process was not subjected to specific examination, and educationists attempting to apply Gestalt hypotheses in detail to classroom practice made deductions which are not necessarily valid. It is true, nevertheless, as Woodworth points out, that Gestalt psychologists believed that associationist psychology, applied to education, supported the traditional emphasis on drill whereas their own theories offered the only hope of children being educated to depend on their intelligence and not on rules. At the close of his career Wertheimer summed up this view as follows: " Repetition is useful, but continuous use of mechanical repetition also has harmful effects. It is dangerous because it easily induces habits of sheer mechanized action, blindness, tendencies to perform slavishly instead of thinking, instead of facing a problem freely."

So far enough has been said to make it clear that Gestalt *theory* rejects synthetic methods and favours a whole-word, meaningful approach to reading. This is only of academic interest, however, unless accompanied by scientific evidence and of this, in the generally accepted sense of the term, there is very little. In much of their work on *human* behaviour, the early, major Gestaltists tended to *demonstrate* phenomena rather than conduct research according to a rigorous experimental design. A few examples of the phenomena used to illustrate some of the main Gestalt principles will, therefore, suffice to show how educationists have often come to accept the work of this school of psychology as ' evidence ' for a whole-word approach to reading.

The perception of a whole before its parts, and the tendency for a whole to dominate and be more than the sum of its parts, was demonstrated in Wertheimer's classic experiments on visual movement. The aim of these studies was to account for the movement seen, for example, in looking at a motion picture, which in objective fact is a series of 'stills' separated by intervals of darkness. Wertheimer's conclusion that this 'Phi-phenomenon' is the result of actual motion in the brain is interesting, but the importance of his work lies in the inference drawn that experience has dynamic wholeness and meaningfulness. Translated in terms of the reading process, this inference implies that a sentence is much more than, and qualitatively different from, the sum of its separate letters, sounds or even words, and will be perceived as such.

If, however, this tendency to perceive wholes before parts operates simply it ought also to be a source of confusion in learning to read. The Gestaltists believed (using figures and geometric forms for their demonstrations) that if, for example, a whole figure appears as circular, slight irregularities are overlooked. By analogy, it would seem that, children would naturally ignore differences in words of similar shape unless specially instructed to detect them.

Another Gestalt law quoted in support of whole-word methods is that of 'proximity'. Dots were effectively used by Wertheimer to demonstrate that they are likely to be seen in groups. Thus, in reading terms, just as . . . . is seen as two groups of five and three dots respectively, so 'brown owl' will be seen as two words and not as eight separate letters.

Gestalt psychologists also played down the role of experience in learning and underlined the importance of insight. They stressed intuitive knowledge and, mainly on the basis of illusion experiments, formed their famous doctrine of the 'innate organisation of perception', which claimed generally that from infancy we perceive the world in the complex, articulated way that we do as adults. From this it would follow that a child's knowledge of two-dimensional form is innate and he will not experience any difficulty in perceiving the pictures and word shapes that he encounters in his first reading book.

Deductions about child development made from the results of inquiries into adult behaviour are naturally suspect, but the difficulty of discovering exactly what a child perceives has led many educationists to accept Gestalt ideas. Before passing on to what is known about *children* which might justify a whole-word approach to reading mention should be made, however, of one piece of evidence which, again by analogy, would appear to throw some light on the difficulties encountered by young children in organising their visual experiences.

Senden (quoted by Hebb) examined the attempts at visual perception of congenitally blind patients who had been successfully operated on for cataract. If his evidence can be accepted, these cases experienced the 'figure-ground' phenomenon of the Gestaltists almost immediately after operation, but they had the utmost difficulty in differentiating between simple shapes such as a square, circle and triangle. Even the development of simple shape perception would seem to require slow and laborious practice, how difficult must it be, therefore, for a young child to discriminate between words, particularly those of similar configuration!

(c) The development of visual perception in young children

The effectiveness of any method for beginning reading, whether it be analytic or synthetic, obviously depends on the readiness of children to make use of it. Factors such as school adjustment, language development, physical and mental maturity, as well as adequate motivation, are very important, but as Dolch points out, " perceptual development is more crucial than any of the others since without perceptual ability, ' telling one word from another ', there is no reading."

Since the accurate perception of shape and pattern, which must logically precede word or letter discrimination, is not very important until a child begins learning to read, few scientific investigations of perceptual development in pre-school children have been carried out. It appears, however, that although children on the average, according to Griffiths, start to look at pictures as early as their second year, their perception may be partial and vague considerably beyond this age. For this reason many authorities insist that children should be able to name pictures without difficulty before beginning reading, and consequently include this type of activity in their reading readiness tests.

More important for reading than picture naming, however, is the ability to perceive two-dimensional shapes without evident representational meaning. How soon this ability develops is doubtful, but there is some evidence that before five years of age the average child will be able to recognise simple, outlined geometrical figures. The more complex figures, as Bender and other investigators have shown, are not fully grasped until later for, although children under five appear to have some awareness of details within a figure, they generally cannot reproduce them. A mental age of seven-plus as the border-line for reasonably accurate form perception has also been reported.

So far the widely-held opinion that young children tend to see shapes as wholes would seem to have some foundation in fact. There are, however, other experiments which show that syncretist perception in children depends on the ' goodness ' of a shape, i.e. in the Gestalt sense of the word, for whenever children have been presented with a relatively unstructured form consisting of salient details the latter have tended to dominate their perceptions. Consequently, further research on this aspect of early form perception needs to be carried out before a final verdict may be given.

In contrast, research workers agree that young children have considerable difficulty with the *orientation* of shapes. Frank, for example, found that the tendency to confuse shapes (represented by playing cards) and their mirror images decreases significantly as children get older, a finding confirmed by Davidson's experiment with silhouette shapes resembling word outlines. The fact that shapes are readily reversed has also been demonstrated by Hildreth and others.

Many psychologists have found, too, that the *sequential order* of shapes presents difficulties for the young child. Teegarden, for instance, asked children who had not yet begun learning to read to name pictures of familiar objects which had been arranged in five rows each of five columns. She reported a singular lack of consistency in the direction of movement used by the children in carrying out this task. This finding has an evident bearing upon the development of skills of word recognition, the aspect of beginning reading to which we now turn.

(d) Cues to word recognition in the early stages of reading

A high correlation between success in the early stages of reading and the ability to perceive and reproduce meaningless shapes might be expected by those who subscribe to the visual pattern—whole-word theory. Various experiments have been carried out to test this hypothesis but none so far has produced a correlation of more than .5. It may be that those, like Kendall, who found no significant relationship, concentrated too much attention on the perceptual abilities of backward readers. A more likely explanation, however, is that too little importance has been attached to the perceptual training and reading instruction previously given to children taking part in these studies. Apparently, in spite of the low correlations reported, a certain minimum perceptual ability is required for reading, but there is evidence that training in the perception of visual form assists children in the beginning and improves their subsequent progress.

Similar difficulties arise in weighing the evidence on how and when children perceive words and letters. Much of the research on this topic has concentrated on the errors children make in word recognition *after* they have been in school for some time. A less vulnerable justification for the widely-held beliefs that analytic methods for beginning reading conform to the 'principles of child development', and that children see whole words before they distinguish their component parts would come from studies of the child's first encounter with words.

An investigation by Ilg and Ames is of this kind. They studied a group of children at each six-monthly age level up to five years and beyond, from the stage of patting an identified picture to silent reading—and reported that it is 'natural' for children under five to recognise and spontaneously demand to print *capital letters*. Other investigators, such as Wilson and Flemming, also tend to confirm that children are letter conscious in the earliest stages of reading.

Daniels and Diack explain the observation of Ilg and Ames on the ground that in children's alphabet books capital letters appear and are more stressed than lower-case letters—an evident environmental rather than genetic factor. This contention may be true, but, while not irrelevant to the main argument, and certainly so if this pre-school environment were under educational control, this preference for the more striking letter patterns at the age of beginning reading cannot be ignored.

Those who support analytic methods for beginning reading frequently refer to a study of pre-school children made by Gates and Boeker in 1923. Gates summarises the results of this early experiment in his book, *The Improvement of Reading*:

> " It was found that beginning pupils observed primarily the length of words and depended upon the observation of the length for later recognition when they were given such a series of words as *cow*, *postman*, *dress*, *duck*, *football* and *dandelion*. To these children differences in length were the most obvious differentiating features. When, however, the words presented at the same time were substantially the same in length the pupils tended to study the words until they found some small but outstanding detail, such as the dot over the *i* in *pig*, the 'funny cross' in *box*, the similar beginning and ending of *window*, and the 'monkey's tail' on the *y* in *monkey*."

From this and other research Gates concludes that, although children seize upon salient details in a word at quite an early stage, dependence upon this method of perceptual attack may lead to difficulty unless other cues to word recognition are developed. These involve synthetic methods such as the use of letters, phonic analysis and syllabification. For beginning reading, however, Gates would seem to approve of an analytic approach since he states that: " In general, the best procedure is to attempt first to recognise the word as a whole."

Some investigators like Frank (cited by Vernon) go so far in their support of whole-word methods as to state categorically that young children perceive the general structure of words and neglect details. Such a finding is in general conformity with the Gestalt hypothesis and the studies of early form perception mentioned earlier. But we must remember, too, that most normal children will have achieved accurate perception of simple forms at a mental age of five-plus. Thus, Frank's observations and those of investigators who claim that children are letter conscious when beginning reading are by no means as contradictory as they appear at first sight; it is at least possible that—whether by environmental pressures or by maturation—the earlier power of gross global perception has been augmented by finer discriminations.

A number of studies have shown that children learn fairly soon to recognise the shapes of certain *isolated* letters. Nevertheless, mainly because letters by themselves are meaningless to a child, advocates of whole-word methods recommend that the shape of single letters should not be taught until a basic sight vocabulary has been acquired. Research findings endorse this recommendation since young children experience considerable difficulty with the orientation of letters with a similar shape. Reversals such as ' b ' for ' d ', ' p ' for ' q ', and inversions like ' b ' and ' p ', ' n ' and ' u ', are common and seem to persist, though with diminished frequency, even with normal readers up to seven or eight years.

It must be remembered that the direction in which letters and words are read is, like the letter symbols themselves, an arbitrary convention. It is not surprising, therefore, that young children are unable to recognise the importance of the order of letters in words. Davidson found that forty-six per cent. of the errors in word matching made by five-year-olds were reversals. Reference to the ' reading gradient ' of Ilg and Ames shows that the transposition of letters within words and complete word reversals, such as ' saw ' for ' was ', persists up till eight or nine years, i.e. longer than letter reversals.

Tendencies to reversals and inversions either of letters or words seem to be at least partly the result of immature perceptual development. Nevertheless, it has been shown, and is indeed self-evident, that some training and practice in the direction in which words should be read—that is in their convention—is essential if normal progress is to be maintained. This is particularly necessary if an analytic approach is adopted, since whole-word methods do not naturally impose any particular habit of directional attack on individual word patterns, although they give greater help than synthetic methods by way of contextual clues. It seems important, therefore, to supplement analytic methods by practice in writing words, since by kinaesthetic experience children will acquire the ' feel ' of a word from left to right.

(e) Reading for meaning from the beginning

Up to this point attention has been focussed mainly upon certain factors in the development of visual perception which might influence the choice of methods for beginning reading. A few inconsistencies in the findings will have been observed, but these may be due to the failure of some investigators to take into full account the influence of the children's experiences prior to the experiment. More research is needed, but, even when the development of visual perception is considered by itself, there appear to be sufficient grounds for the belief that initially an analytic approach to reading is likely to be more suitable than a synthetic one.

Reading, however, is much more than the mechanical recognition of words; it is a thought-getting process involving meaning and purpose and as such a logical development of language. It follows, therefore, that in deciding the appropriate method for introducing children to reading, careful attention must be paid to the role that spoken and written language plays not only in the reading lesson itself but in any present or future learning situation.

In *The Language and Mental Development of Children* by Watts we are told that “ the average child enters the public elementary school at five plus with a vocabulary of at least 2,000 words (ample for any educational adventure with an intelligent teacher).” These words are ‘ wholes ’ with meaningful associations which may be used either singly, in a phrase, or in a sentence, to express ideas—the natural units of thought. Since whole words and not letters, sounds, or syllables are the natural units of linguistic expression it seems logical that whole-word methods should be used to introduce children to reading.

Even those who consider that not enough emphasis is placed upon synthetic methods, agree that since children's speech consists of whole words, early reading ought also to be based on whole words. The new *phonic-word-method* of Daniels and Diack, resembles analytic methods in “ its insistence upon reading for meaning and in not teaching the sound values of letters in isolation from whole words.”

Previously, it was generally asserted that whole-word methods develop the ability to read with speed, fluency and understanding, but fail to encourage the accurate word recognition and correct spelling which follow from a synthetic approach. There is, naturally, very little proof of the soundness of this contention since it has been shown that most teachers supplement whole-word methods by synthetic methods at some stage.

Visits to over a hundred primary schools in connection with the Foundation's recent inquiries suggested that classrooms where materials based upon a whole-word *introduction* to reading were being used were generally more attractive and stimulating than those few in which the accent was on the acquisition of phonic skill. Although this kind of observation may be considered ‘ unscientific ’, the fact that it has been repeated over and over again by school inspectors and the like would suggest that its validation could be obtained.

Summary and Conclusions

Arguments for and against whole-word methods of teaching reading become so heated that it is easy to lose sight of the fact that the main problem in the first years is that of the correct timing of the various approaches. Children under five years of age generally find difficulty in perceiving small details in shapes. Normal pupils in the reception classes of our schools will have developed the ability to perceive simple figures, but will not be able to remember and reproduce accurately more complicated forms such as words, although they may be aware of their details. Accurate form perception, like phonic readiness, is apparently not achieved by most children under normal educational conditions in this country, until a mental age of seven. This means that some children even in the upper junior school will still have difficulty with the orientation and sequential order of letters in words. Conversely, the abler children—and particularly those whose early years have been rich in pre-reading experiences—will achieve perceptual maturity much sooner and in some cases by the time they start school. Thus, a synthetic approach to reading with its emphasis on small, meaningless parts of words is likely to be more difficult for most young children than a more analytic, word-whole one.

There is no suggestion that children can learn to read by word, phrase, sentence and story methods alone. Even when these methods are used in the early stages, as would seem appropriate, there is need for plenty of writing experiences to emphasise the left-to-right directional attack upon words. Synthetic methods, too, are necessary at a later stage to develop greater accuracy in word recognition, and to give children a means of tackling new words independently.

Children in the reception class will be at various stages of reading readiness requiring different methods of teaching. The majority will need to start with a whole-word approach followed later by phonic instruction based upon the vocabulary already acquired and understood; a few, because of high intelligence and a favourable home background, will be beyond the introductory, analytic stage, and ready for phonic training and word building immediately.

Readiness, which involves a certain perceptual, intellectual, emotional and social maturity in children, is clearly the key to the successful use of both analytic and synthetic methods of teaching reading. In deciding the best method or methods for children at any stage of development it is important, however, never to forget that the purpose of reading is to obtain *meaning* from printed symbols. Whole-word methods for beginning reading, properly applied, achieve this object naturally, but unless used with care, synthetic methods may fail to do so. The end result is of such vast educational significance that this must be the primary justification for any method.

Bibliography

BENDER, L., *A Visual Motor Gestalt Test and its Clinical Use.* New York: American Orthopsychiatric Association, 1938.
BOARD OF EDUCATION, *Infant and Nursery Schools.* London: H.M. Stationery Office, 1933.
BOYCE, E. R., *Learning to Read: A Handbook for Teachers and Parents.* London: Macmillan, 1950.
BOYCE, E. R., *The Improvement of Reading in the Junior School.* Leeds: E. J. Arnold & Son (n.d.).
CATTELL, J. McK., " Ueber die Zeit der Erkennung und Benennung von Schriftzeichen, Bildern und Farben." *Philos. St.*, 2, 1885. Quoted by Woodworth, R. S., and Schlosberg, H., *Experimental Psychology.* London: Methuen, 1954.
DANIELS, J. C., and DIACK, H., *Learning to Read.* London: News Chronicle, 1954.
DANIELS, J. C., and DIACK, H., *Progress in Reading.* University of Nottingham Institute of Education, 1956.
DAVIDSON, H. P., " A Study of Reversals in Young Children." *Journal of Genetic Psychology*, Vol. XLV, 1934.
DIACK, H., " First Steps in Reading: Phonics the Key." *Times Educational Supplement*, 7th May, 1954.
DOLCH, E. W., *Psychology and Teaching of Reading.* Champaign, Illinois: Garrard Press, 1951.
DUNCAN, J., *Backwardness in Reading.* London: Harrap, 1953.

Ellis, W. D., *A Source Book of Gestalt Psychology*. New York: Harcourt, Brace & Co., 1938.

Erdmann, B., and Dodge, R., *Psychologische Untersuchungen über das Lesen*. Halle: Niemeyer, 1898. Quoted by Woodworth, R. S., and Schlosberg, H., *Experimental Psychology*. London: Methuen, 1954.

Frank, H., " A Comparative Study of Children who are Backward in Reading and Beginners in the Infant School." *Brit. J. Ed. Psych.*, Vol. V, 1935.

Gagg, J. C., " First Steps in Reading: Present Practice." *Times Educational Supplement*, 14th May, 1954.

Gates, A. I., *The Improvement of Reading*. New York: Macmillan, 3rd ed., 1947.

Gates, A. I., and Boeker, E., " A Study of Initial Stages in Reading by Pre-school Children." *Teachers College Record*, November, 1923.

Gray, W. S., *On Their Own in Reading*. Chicago: Scott, Foresman & Co., 1948.

Griffiths, R., *The Abilities of Babies*. University of London Press, 1954.

Hebb, D. O., *The Organisation of Behaviour*. New York: Wiley, 1949.

Hildreth, G., " Reversals in Reading and Writing." *J. Ed. Psych.*, Vol. XXV, 1934.

Ilg, F. L., and Ames, L., "Developmental Trends in Reading Behaviour." *Journal of Genetic Psychology*, Vol. LXXVI, 1950.

Kendall, B. S., " A Note on the Relation of Retardation in Reading to Performance on a Memory-for-Designs Test." *J. Ed. Psych.*, Vol. XXXIX, October, 1948.

Ministry of Education, *Reading Ability: Some Suggestions for Helping the Backward*. London: H.M. Stationery Office, 1950.

Ministry of Education, *Language: Some Suggestions for Teachers of English and Others*. London: H.M. Stationery Office, 1954.

Ministry of Education, *Standards of Reading*, 1948 *to* 1956. London: H.M. Stationery Office, 1957.

Monroe, P., *A Brief Course in the History of Education*. New York: Macmillan, 1909.

Morris, J. M., " Teaching Children to Read, I. The Relative Effectiveness of Different Methods of Teaching Reading. A.—The Place and Value of Phonics." *Educational Research*, November, 1958.

Morris, J. M., *Reading in the Primary School*. Publication No. 12 of the National Foundation for Educational Research. London: Newnes, 1959.

Russell, D. H., " Teachers' Views on Phonics." *Elementary English*, Vol. XXXII, October, 1955.

Senden, M. v., *Raum—und Gestaltauffassung bei operierten Blindgeborenen vor und nach der Operation*. Leipzig: Barth, 1932. Quoted by Hebb, D. O., *The Organisation of Behaviour*. New York: Wiley, 1949.

Smith, N. B., *American Reading Instruction*. New York: Silver, Burdett and Co., 1934.

Teegarden, L., " Tests for the Tendency to Reversal in Reading." *J. Educ. Res.*, Vol. XXVII, October, 1933.

Unesco I.B.E. Twelfth International Conference on Public Education. *The Teaching of Reading*. Geneva, 1949.

Vernon, M. D., *Visual Perception*. Cambridge University Press, 1937.

Vernon, M. D., *A Further Study of Visual Perception*. Cambridge University Press, 1954.

Vernon, M. D., *Backwardness in Reading*. Cambridge University Press, 1957.

Watts, A. F., *The Language and Mental Development of Children*. London: Harrap, 1944.

Wertheimer, M., *Productive Thinking*. New York: Harper, 1945.

Whitehead, F., " Rival Reading Methods: Question of Timing." *Times Educational Supplement*, 21st May, 1954.

Wilson, F. T., and Fleming, C. W., " Letter Consciousness of Beginners in Reading." *Journal of Genetic Psychology*, Vol. LIII, 1938.

Woodworth, R. S., *Contemporary Schools of Psychology*. London: Methuen, 1949.

Woodworth, R. S., and Schlosberg, H., *Experimental Psychology*. London: Methuen, 1954.

A STUDY OF VOCABULARY ACHIEVEMENT IN THE ENGLISH LANGUAGE IN SECONDARY SCHOOL CHILDREN

Thesis submitted for Ph.D., Leeds, 1958.

by D. G. BURNS

Aim

The principal purpose of the enquiry was to classify a list of words using children's definitions as a criterion. The definitions were supplied by senior pupils in secondary modern schools and it was hoped that the classified list could be used as a standard of reference for younger pupils in the same type of school and for young adults (over 15 years of age) who had recently left school.

Method

The enquiry fell into two parts: (*a*) the preparation of a list of words which children in secondary modern schools were likely to meet in reading; and (*b*) an examination of the definitions of these words given by pupils in their third year of study.

(*a*) It was assumed that children in their third year of study in secondary modern schools (ages 13 and 14) might be able to give definitions of some 9,000-10,000 words and a list of this size was compiled from various sources. The principal source was *The Teachers' Word Book* and all words in the frequency ranges from AA to 5 (roughly 9,200 in all) were provisionally included. A separate enquiry was made to discover the books, newspapers and other periodicals read by children of this age attending secondary modern schools in an urban area and a list was made of the words occurring in samples taken from these sources which did not occur in the selected portions of *The Teachers' Word Book*. This provided some 750 additional words. It was evident that many of the words in this provisional list would be known to pupils when they entered the secondary modern school at 11-plus and, in fact, investigation in the schools revealed that nearly all pupils were familiar with some 4,800 of them. There were also a number of words taken from *The Teachers' Word Book* (archaisms, abbreviations, foreign words and—words unlikely to be familiar to British children like *campus*, *chipmunk* and *levée*) which did not seem relevant to the enquiry. These and the words known to most children at 11-plus were deleted from the provisional list, and the remainder (some 4,700 words in all) were classified using the children's definitions as a criterion.

(*b*) Arrangements were made by courtesy of the Chief Education Officer in Leeds to examine all the pupils in twenty large secondary modern schools who were in their third year of study (3,684 pupils participated), representing about 80 per cent. of all pupils in the senior classes of non-selective secondary schools in the area. The pupils lived in all parts of the urban area and the group was representative of all children attending the secondary modern schools of the L.E.A. The children in this group had I.Q.s ranging from 70 to 115 (according to a Moray House test administered at 10-plus) and since it was evident that the wide range of ability and attainments in such a group would be reflected in their definitions of words, it was important to ensure that the definitions obtained for each word were provided either by the whole group of pupils or a representative sample of them. It was not possible to examine many pupils on such a large number of words but precautions were taken to ensure that the definitions received for each word were representative of the whole group. The list was divided into fifty sections of about 90 words each and scripts were prepared to cover the words in each of these sections. Copies of the scripts were interleaved and were distributed in the schools serially so that each school received a proportionate number of each of the scripts and the same method of distribution was used in each class within the school. In this way, the scripts completed for each section of the list were answered by a representative group of pupils in all the classes of the participating schools. A minimum of 200 definitions was obtained for each word in the list, approximately equal numbers of boys and girls participated in the enquiry and separate records were kept for each sex.

As a number of the words in the list have more than one meaning, a decision had to be taken whether to examine children's knowledge of several meanings of the same word or only one. The number of words was so large that it did not seem possible to adopt the first alternative and the enquiry was limited accordingly to examining pupils' knowledge of one meaning of each word in the list. Each of the words in the list was presented in a brief context, the simplest way of showing the word in use, the word in the context was underlined and pupils were asked to write their definitions of it on the scripts provided. The definitions were marked according to a three-point scale (1, $\frac{1}{2}$ or 0) which provided a numerical score for each word in the list.

Such a method of enquiry has evident limitations. It assumes, for example, that meaning can be defined although this remains the central problem of semantic studies and that a scale of values exists by which definitions may be compared. There is no general agreement on the theoretical premises which might justify such assumptions and though criteria were formulated in this enquiry to help in distinguishing more consistently between good definitions and bad, the interpretation of what is good or bad is ultimately a subjective one. It has been argued, for example, that definitions which appeared to be logically and operationally incomplete could only be partially satisfactory (A *cask* of wine . . . wooden container: The *caliph* of Baghdad . . . prince): and on the same grounds, that definitions which placed a word correctly in its semantic family but omitted characteristics which would distinguish it from other words in the same family were also only partially satisfactory (He *cantered* into the yard . . . galloped: A *cleft* in the rock . . . opening). Such criteria are arbitrary and to this extent, the classification of words based upon them, must be considered a subjective one.

Secondly, the gradings which have been given to the words in the list reveal children's knowledge of them in a particular context and may give no more than an approximate indication of how well they would be understood in a different one.

Thirdly, the scope of the investigation was limited to a selection of some 4,700 of the words which children are likely to meet in reading and though this is a representative sample of the words used in the sources consulted, it is not in any way representative of all the other words which children will probably meet in reading.

D.G.B.

In spite of these limitations, the present list, which is believed to be the only one of its kind based on the performance of British children, will be of particular help to research workers and others engaged in the preparation or assessment of reading materials for pupils attending secondary modern schools. It may also be of use in the preparation of tests of attainment in reading and in estimating the difficulty of reading materials.

W.D.W.

Bulk Orders for off-prints of articles appearing in "Educational Research"

We have received requests from Training Colleges and Departments of Education for considerable numbers of off-prints of certain articles for the use of students. This is a service which we shall be pleased to give but it is impracticable to hold large stocks of off-prints on the chance that they may be called for. If, however, the Editor is notified in time that considerable numbers of off-prints are required, it will be possible to print these specially in a handy form and at a reasonable price. A request for a quotation for a bulk order of off-prints must be received within one month of the publication of the issue in which the article appears, and should be addressed to the Editor, "Educational Research," 79, Wimpole Street, London, W.1.

OCCASIONAL PUBLICATIONS

THE Foundation is beginning a small series of ' Occasional Publications '. These will be documents issued in duplicated or off-set lithograph form suitably bound and priced to cover the costs of materials, and postage in the United Kingdom.

The series is intended to make available to research workers and others, specialised matter arising from research work undertaken in this country by the N.F.E.R. itself, by individual research workers, by officers attached to L.E.A.'s or by the research departments of Universities. Material which would normally find its place in one or other of the learned journals will not be published, nor will matter which could be expected to achieve publication in book form. The series is intended to provide an outlet for work of a high level but with a limited readership and which would otherwise become rapidly inaccessible.

Each manuscript for inclusion in the series will be considered on its merits by the Foundation in the light of the validity of the research on which it is based, its relevance to some aspect of the Foundation's own programme and to educational research in this country, the possibilities of publication elsewhere and the value of the information it contains to the education system or to individual workers. It is not anticipated that more than three or four such documents will be issued in each year.

Vocabulary of Secondary Modern Schools

The first of this series will be " The Vocabulary of Secondary Modern School Children," by D. G. Burns. It will consist of the word frequency lists arrived at by Dr. Burns as the result of his research in Leeds preceded by a brief account of the method by which they were arrived at and some of the results of his study. The document should be of importance to writers of textbooks for the Secondary Modern school, those concerned with remedial reading and those who wish to construct tests of vocabulary or comprehension.

The cost of Occasional Paper No. 1 will be approximately 8/6, post free in U.K. It will be published in May, 1959, and orders should be sent to the National Foundation for Educational Research, 79, Wimpole Street, London, W.1.

The Foundation will be pleased to receive from members and from individuals suggestions for material for inclusion in the series. It is advisable in the first instance to write to the Director of the N.F.E.R. indicating the nature of the material proposed for inclusion.

BINDER FOR EDUCATIONAL RESEARCH

A binder has been produced to hold three volumes (nine issues) and an index of *Educational Research*. It is in an attractive colour with the issue and volume numbers in gold on the spine. The binder may be obtained from the publishers, Newnes Educational Publishing Co., Tower House, Southampton Street, Strand, London, W.C.2, at a price of 10/6.

TEST ANNOUNCEMENTS

Recent additions to the Foundation's "open" series of tests:

Mechanical Arithmetic Tests 2A and 2B. These are parallel tests devised to measure attainment in mechanical arithmetic of children in the first year of the primary school. They are fully standardised over the age range 7 : 0 to 8 : 8.

Arithmetic Progress Tests B1 and B2. These are parallel tests intended to provide an estimate of the amount of the arithmetic curriculum assimilated by children in the third year of the primary school, and are fully standardised over the age range 9 : 0 to 10 : 8.

Further details of these and the other "open" tests prepared for the use of the teacher in the classroom are given in the catalogue, *Educational Guidance in Schools: Standardised Tests for the Use of Teachers*, obtainable from the Foundation.

Test Agency

One of the aims of the Agency is to make available to users in this country the many tests originating overseas that, because of import restrictions, are difficult to obtain. In pursuance of this policy the Foundation has the agency for the distribution of tests and test materials from the Australian Council for Educational Research; the Psychological Corporation, New York; Science Research Associates, Chicago; the California Test Bureau; and many other publishers.

In addition, the Test Agency distributes in the United Kingdom and overseas a number of psychological and educational tests devised by British psychologists.

Full details and prices of all tests distributed by the Agency will be found in the new edition of the Test Agency catalogue shortly available from the Foundation.

The following new test has recently been made available:

The "STYCAR" Hearing Tests, by M. D. SHERIDAN, M.A., M.D., D.C.H.

These consist of a series of simple clinical auditory screening tests which have been found useful in the preliminary assessment of the everyday hearing of very young or mentally handicapped children. They are designed to obtain reliable information concerning a child's capacity to hear with comprehension in everyday situations. These tests are restricted in sale to qualified psychologists and medical practitioners, and are supplied with a fully descriptive manual. Further details and prices are available from the Foundation.

Test Library

As part of its service to test-users, the Foundation has built up a comprehensive library of tests and test materials. This library is available for reference at the Foundation and where possible advice and guidance on the use of tests will be given.

© *National Foundation for Educational Research in England and Wales.*

Printed by King, Thorne & Stace Limited, 1-13 Jubilee Street, Brighton, 1,
and published by Newnes Educational Publishing Co., Ltd., Tower House, Southampton Street, Strand, London, W.C.2.,
for The National Foundation for Educational Research in England and Wales, 79 Wimpole Street, London, W.1.

VOLUME 1. NO. 3 JUNE, 1959

EDUCATIONAL RESEARCH

A Review for
Teachers and all concerned with
Progress in Education

Published for
THE NATIONAL FOUNDATION FOR EDUCATIONAL
RESEARCH IN ENGLAND AND WALES

by

NEWNES EDUCATIONAL PUBLISHING CO. LTD.
Tower House, Southampton Street, Strand, London, W.C.2

in each case. It would be interesting to know what basis there was, other than that of increasing specialisation higher up the age range, for making the minimum class size higher for the primary than for the secondary school. It may well be that the average child will be served best by some experience of large classes, some experience of small classes, and some individual tuition, all these intelligently related to a considerable variety of factors: available space, the different nature of different subjects and activities, the different stages of development and the different personalities of the children in the school."

In her reply Dr. Fleming points out that our correspondent has warned us of the complexity of the issue—a complexity shared by almost any topic of the social and educational sciences—a warning she intended to convey by the wording of the title and by the brief survey at the end of some of the other variables in the teaching situation. She continues, " Popular opinion has often been appealed to in the past in support of action based on the belief that the quality of teachers is of less significance than their quantity. The profession has been diluted by the admission of untrained personnel. Salaries of highly skilled teachers have remained disproportionately low and adequate provision has not been made of the secretarial assistance, classroom space, books and material equipment which lighten a teacher's load. Much of this has been justified explicitly or implicitly under cover of the plea that smaller classes are a top priority; and the emphasis on size of class has tended to deflect attention from the conditions actually contributory to high morale and professional pride among teachers. It is partly for this reason that I am quite without regret as to my decision to provide a survey of the enquiries which (to the extent indicated) challenge the definitive character of mere size. At the same time I point again to the accumulating and influential evidence in support of the significance of skilled teaching in wholesome social climates. . . . As one convinced of the importance of due humility in face of ascertainable facts I submit for careful scrutiny the challenge contained in the many pointers past the single issue of size of group to the character of the relationships and the quality of the personalities involved in the processes of education."

This answers some of the points made by our critic; but it doesn't answer all of them. It should be said that few of the articles in this journal are intended to be complete in themselves, since most educational issues are too complex for a full treatment of all their aspects in 7,000 words. Authors are asked to present the evidence on one aspect, and we expect our readers to appreciate this, and to look forward to further articles on other aspects. We hope, for example, to have a paper on the structure of social relationships which develops in class groups of different sizes.

What, then, of the other issues? Much research in education suffers because the wrong or an irrelevant question is put; but educational research is not exceptional in this and the natural sciences are strewn with equally glaring examples. In the present case is the question irrelevant or wrongly put? One aspect of the problem is surely the question of relationship between size of class (within the range of sizes as we, at present, find them) and the formal attainments of children. To say, because we find no such relationship apparent, that we should look for other " intangible " effects or that we should experiment with smaller, larger or different types of class, is to pose other, and possibly more important questions, but not more valid ones. Such questions are the next stage of an enquiry on class size. To find objective answers to them requires at least two things: the willingness and the ability of schools to experiment for a considerable length of time with groups of many different sizes for pupils of different ages and to relate these variations in size to different definable educational purposes; and the money to conduct the enquiry. On educational research in this country is spent annually in all (at a generous estimate) less than £100,000, or approximately .017 per cent. of the total national expenditure on education.

More apparently damaging is the criticism that, if the questions are correctly posed they involve " intangibles " not susceptible of objective study or measurement—and that anyway the educational sciences are not sciences in the rigorous sense of the natural sciences. This is a frequently heard criticism and has a certain justification. But those who make it forget a number of things. The infancy of a science is often a period of trial and error in which new weapons have to be forged and tested—and many abandoned. Moreover, the older sciences possess, by the nature of their material, the possibility of reduplication of experiences, of expenditure of material and of direct transfer from artificial laboratory conditions where factors can be varied independently to " real " situations. The human sciences are rarely like this; and their methods are necessarily different—certainly where the psychological development of children is concerned. Rigorous, objective thought however should be common to both; so, too, is the verification of hypotheses by testing their predictive value. The so-called " intangibles "—if they are susceptible of definition and if they are not they are merely a cover for imprecise thought—must be susceptible of some kind of measurement and it must be possible either to predict them or to use them as predictors. That this has not been satisfactorily done for all the intangibles or even for many (though it has for some) is because agreed definitions are lacking and because there has not been the massive devotion of time and money necessary to the attempt to measure them.

May we, in conclusion, come back to the point of the responsibility of a journal such as this. Of course we expect our colleagues to regard critically such findings as are put forward and to test them in the light of their own experience. Of course policy decisions which affect human beings and the very nature of our society cannot be mechanically taken and policy-making in any field is always an " art " and not a " science." The difficulty and challenge of work in educational psychology is that there can never be one right answer. What can be done is a painstaking exploration of the consequences to children of different decisions which are made and to present these as the basis for varying or ratifying them. All education implies choice; and choice implies sacrifice as well as gain. Research can cast something of the profit and loss account. Dr. Fleming's article said in effect that by reducing the size of classes within the now possible range and, *doing that alone*, your profits as expressed in attainment, will not go up. There may well be other advantages and disadvantages but they must be sought elsewhere.

SUBSCRIPTION RENEWAL FORM

To
NEWNES EDUCATIONAL PUBLISHING CO., LTD.
TOWER HOUSE, SOUTHAMPTON STREET,
LONDON, W.C.2.

Dear Sirs,

Please supply .... copy/copies of EDUCATIONAL RESEARCH VOL. 11, Nos. 4, 5 & 6. I enclose cheque/money order/postal order for £.......... ....s. ....d. at rate of 5/6 (post free) per copy or 16/6 (post free) for Annual Subscription covering one copy each of three issues.

NAME..
ADDRESS .... BLOCK
.... CAPITALS
.... PLEASE

To save unnecessary postage it will be appreciated if the remittance is sent with the Order Form.

....
Signature

BULK ORDERS FOR OFF-PRINTS OF ARTICLES

We have received requests from Training Colleges and Departments of Education for numbers of off-prints of certain articles for the use of students. This is a service which we shall be pleased to give but it is impracticable to hold large stocks of off-prints on the chance that they may be called for. If, however, the Editor is notified in time that considerable numbers of off-prints are required, it will be possible to print these specially in a handy form and at a reasonable price. A request for a quotation for a bulk order of off-prints must be received within one month of the publication of the issue in which the article appears, and should be addressed to the Editor, "Educational Research," 79, Wimpole Street, London, W.1.

SPECIAL BINDER

A binder has been produced to hold three volumes (nine issues) and an index of *Educational Research*. It is in an attractive colour with the issue and volume numbers in gold on the spine. The binder may be obtained from the publishers, Newnes Educational Publishing Co., Tower House, Southampton Street, Strand, London, W.C.2, at a price of 10/6 (plus postage 1/6).

The Teaching of Mathematics—II

ATTITUDES TO ARITHMETIC–NUMBER ANXIETY

by J. B. Biggs

Introduction

IT is becoming increasingly recognised by the general public that probably the greatest problem in the teaching of mathematics is to overcome the fear or anxiety the subject is popularly believed to instil in the mind of the learner. The recent pamphlet issued by the Ministry of Education, *The Teaching of Mathematics in Secondary Schools*, had many things to say about mathematics teaching, but it is most significant that the Press almost to a man to commented upon this one aspect—" Distasteful Maths " (*Daily Telegraph*), " Fear or Fun in Maths? " (*Observer*), " How to Frighten People Off Maths " (*News Chronicle*), " BLACK MARKS for those DREARY OLD SUMS " (*Daily Mirror*). Asked to comment upon the pamphlet, Professor J. Lauwerys, in the B.B.C. *At Home and Abroad* programme, stated: " No subject arouses so much distaste in boys and girls as mathematics . . . Girls definitely suffer more from this distaste than do boys."

Some months prior to this, there was considerable discussion about the shortage of scientists and technicians in relation to the Ministry of Education's proposals to increase the teacher-training course from two to three years. In a leading article (" A Rush of Teachers," September 25th, 1958), *The Times* suggested that a cause of this shortage, and especially of scientists at the top, might lie in " the deplorable quality of the mathematics teaching in many primary schools, which has equipped children with an original distaste for the subject as strong as Gibbon's. . . ."

It is unnecessary to multiply examples of the pre-occupation with this aspect of mathematics teaching. It is evident that here is a problem as great as any in education, but one where opinion flourishes in proportion to the lack of definitive research findings. There is, however, some work which is relevant and suggestive, and which it is worthwhile to summarise, if only to clear the way for further, more conclusive, research.

The " Emotional Component " In Mathematical Ability

Factor analytic studies into mathematical ability have shown that there appear to be two main types of factors which account for the greater part of variance in the obtained test scores: these are the cognitive and the affective or emotional factors.

The studies agree that the one cognitive variable most predictive of success in mathematics is " g," or general intelligence (Wrigley, 1958). Others, such as " number," " spatial " and " verbal " factors, have also been shown to be operative. This kind of research naturally only identifies these factors—it does not go further into the explanation of their nature. It is possible that these subsidiary factors are genetic in origin: a simpler and more likely explanation, however, is that they are the result of teaching and number experiences generally. Thus, research subsequent to Piaget's own investigations has shown that the " stages " in number development which he describes are dependent upon two things—mental age (or "g" in the present terminology), and kinds of experience during the formative periods of number concept development. Some of these cognitive aspects have already been discussed in the first article of this series.*

This leaves the affective aspects of mathematical ability, or the "emotional component," as some investigators call it. This may also be complex and made up of different kinds of factors. On the one hand, it seems that certain types of personality are better suited to mathematical study than others: emotional instability and anxiety have been shown to affect adversely performance in mathematics, as that subject is at present taught, to a relatively greater extent than performance in other subjects. This aspect may be called a *primary* factor of emotionality in mathematics, as opposed to the secondary or derived factors which are specific to mathematics and which are not necessarily part of the basic personality pattern. These secondary factors are, of course, the attitudes displayed by children and adults towards arithmetic and mathematics—the negative ones varying from a simple distaste to what amounts to a veritable phobia. It scarcely needs pointing out, however, that these attitudes may tend to take root more easily in the unstable personality.

There would appear to be no reason why an intelligent child should not become adept at mathematics *if:* (*a*) he is taught in a logically and psychologically sound manner from the very first stages; and (*b*) he is emotionally stable. Furthermore, if these conditions are fulfilled, it is difficult to see why the much-vaunted fear of number should make an appearance at all. There is no scientific ground for assuming that there are Word People and Number People, as one correspondent in *The Observer* ingenuousy suggested, and even less that some of us might have a "Fear of Number born in us."† There has been some work reported in the psychoanalytic literature, by, for example, Wegrocki (1938) and Bartemeier (1941), who describe interesting case histories of phobias towards specific numbers. Jeffreys (1936) and the Planks (1954) enunciate some general principles concerning the dynamics of numbers and numerical processes in psychoanalytic terms, but these are rather tenuous when viewed in the light of classroom practice and accordingly no more than a passing reference can be afforded them here.

However, as was noted in the previous article in this series, mathematical teaching in the primary school is at best only roughly correlated with what we know about the intellectual development of the child, and therefore many do not achieve optimal understanding of the subject. If this is serious at any stage, negative attitudes and anxieties directed towards mathematics may well arise—based originally on sheer puzzlement. It is intended in this article to look into the research findings related to emotion in mathematics, and in particular to consider their implications for the teacher.

Is Mathematics Intrinsically Distasteful?

In view of the awareness, amongst the general public, of a widespread distaste for mathematics, this question is particularly pertinent. In order to answer it, it is first necessary to enumerate briefly those features which are peculiar to mathematics as a school subject. The first which springs to mind is its *abstract* nature. Mathematics does demand even from the outset a level of abstraction higher than that for most other subjects—and this is in itself a black mark in the eyes of many people to whom the process of abstraction is by no means easy. As we saw in the previous article, the Piaget-Inhelder experiments demonstrated that the capacity for true abstraction is not attained until adolescence and then is developed in varying degrees—in some people, not at all. However, it is present in most, and the popularity of certain kinds of games and puzzles in the Sunday press suggests that many people enjoy making use of this ability. It is relevant here to note that Mace suggests that " mathematical phobia," as he calls it, may be due to the fact that certain branches of mathematics are often taught long before the required powers of abstraction have matured. This, however, is a fault of teaching method, and is not itself due to the abstract nature of mathematics.

---

* *Educational Research*, Vol. I, No. 2.

† *The Observer*, February 1, 1959.

A second feature of mathematics, not shared by the literary subjects, is that the answers to the vast majority of mathematical problems are either right or wrong—one may not get full or even good marks for " elegant style " alone, at least not at the level of mathematics with which we are concerned. This " either-or " nature of mathematics might well prove disturbing and distasteful to certain anxious personalities who do not like to commit themselves in any situation, let alone the mathematical. This diffuse anxiety reaction, however, is an adult kind of defence, and should not apply to young children to whom a specific fear of failure, based upon past experience of failure in a definite situation, is of much greater importance.

Prima facie, there is every suggestion that the " nature " of mathematics should be such as to appeal to any healthy, enquiring mind. Russell pleads eloquently for this appeal in his book, *Mysticism and Logic:* " Mathematics, rightly viewed, possesses not only truth, but supreme beauty—a beauty cold and austere. . . . The true spirit of delight, the exaltation, the sense of being more than Man, which is the touchstone of the highest excellence, is to be found in mathematics as surely as in poetry."

" Rightly viewed " seem to be the operative words here, and the fact that so many people do *not* sense being more than Man while studying mathematics, Russell considers to be the fault of those " dry pedants " who bid the eager youth who would press forward to be initiated to the domes and arches of the sacred temple of mathematics, to " turn back and count the steps."

But probably more convincing than Russell's rhetoric is the amount of empirical evidence available to show that a substantial number of pupils of both sexes do not necessarily dislike arithmetic. One study performed in California (Dutton, 1956) with 459 junior high school pupils (slightly older than U.K. children at secondary transfer) showed that 68 per cent. of the children rated themselves, on an 11-point scale, as " liking " the subject, 15 per cent. were neutral, and only 16 per cent. expressed an active dislike. It should be pointed out, too, that these data were gathered anonymously; thus, presumably, the children gave franker replies. Only 5 per cent. of the children in response to the questionnaire said that they " detest arithmetic and avoid using it at all times," whereas at the other extreme 10 per cent. said that " arithmetic thrills me, and I like it better than any other subject." Most pupils (87 per cent.) enjoy problems when they can do them well, and 83 per cent. feel that arithmetic is as important as any other subject.

In this country, Freeman (1948) administered a questionnaire, again anonymously, to 2,000 pupils from all types of secondary schools. About half said that they liked arithmetic " fairly well " and problems " sometimes." The pattern of responses to this part of the questionnaire was similar for all types of school organisation. More reference will be given to this important study later.

Wall (1945) reports that when a group of young industrial workers were asked to indicate what subjects they would like to take in a voluntary day school, 69 per cent. of the boys and 32 per cent. of the girls* signified that they would like to participate in an arithmetic class. Such a result is scarcely in keeping with the notion that the majority of these un- and semi-skilled youths found mathematics very distasteful. In a pilot study in the N.F.E.R. research programme, 139 third and fourth year junior school children were given a scale consisting of eleven items, and were asked to check the item which most nearly described their own feelings about arithmetic. Forty-five per cent checked the neutral item (" I am not really keen on arithmetic, but I do not dislike it either "), 35 per cent. checked items which were favourable to arithmetic (from " quite nice " to " really thrills me "), and only 20 per cent. registered unfavourable attitudes (from " don't really enjoy " to " absolutely loathe "). On the other hand, in another school, attitudes were measured by means of a sentence completion test and here it was found that 25 per cent. of the responses (each child giving several) were favourable to arithmetic and 52 per cent. favourable to English; 53 per cent. and 35 per cent. were unfavourable to arithmetic and English respectively; and the remaining responses indicated indifference.

We may conclude, then, that mathematics and arithmetic are not intrinsically distasteful, although dislike for these subjects may be more readily aroused than for other school subjects.

Origins and Causes of Attitudes

Before entering into an involved discussion about the relationship of the primary emotional factors to mathematical performance, we should first look at the more obvious secondary factors, i.e. negative (and positive) attitudes specifically directed towards arithmetic. Dutton (1956), whom we have already mentioned, claims that most of the children he tested felt that they knew when their feelings towards arithmetic were developed. Grades 3 to 8, and especially grades 5 and 7 (i.e. approximately at ages 10 and 12), were most significant for attitude formation. However, it must be remembered that this information was obtained on the basis of the pupil's own memory—which is not likely to be entirely reliable, especially concerning the earliest school years. Dutton also asked them if their attitudes had changed at all in the course of their school careers, and about one-third of the children claimed that they probably had done so. These data collectively would suggest that attitudes as measured by Dutton are very much a function of the immediate syllabus and may not be particularly deep-rooted.

Churchill (1958) seems to be rather nearer the mark when she suggests that faulty development of number concepts is probably the most likely cause of the " strong dislike and even fear " which many intelligent adults are known to manifest towards arithmetical operations. This conception places the first origins of the phenomenon as far back as the *infant* school, where critics of current practice assert that while children are taught perhaps to calculate, their number understanding is rarely developed. If the young child does not grasp the significance of number relations at this stage, subsequent arithmetic lessons can only be meaningless and puzzling—although he may mislead both his teachers and himself by obtaining " correct " answers which, however, are the result of slavishly following a mumbo-jumbo of rules. Williams (1958) also agrees with this view when he says that a disparity between conceptual and mechanical attainments leads to a dislike of mathematics. This, too, is one of the *raisons d'être* of certain methods of teaching arithmetic: Stern, for example, claims in her book that, as her method " teaches for understanding," it can be presented in kindergarten (4-7 years) to the " joy of children, teachers and parents."

---

*Little reference will be made to sex differences in attitudes to arithmetic throughout the present article since this issue is tied up very much with the wider one of sex differences generally in mathematical performance. This whole problem will be dealt with in a subsequent article.

Related to the question of the origin of attitudes to arithmetic, is the more important one of the causes of these attitudes. Some of these are self-evident. Snell, in a presidential address to the London Branch of the Mathematical Association, is reported in *T.E.S.** as saying: " The ' don't-like ' attitude of some scholars towards mathematics was often the result of trying too much too quickly, with resultant disillusionment and failure." We have seen that Mace's opinion is that this same phenomenon is due to too much abstraction too early; the authors of the Ministry pamphlet are concerned about the excessively formal nature of much of the mathematical work forced upon children at the present time, and with poor attitudes and lack of true mathematical understanding on the part of the teachers—themselves the products of the same system. Drummond writes that sensitive, intelligent children, if they are hurried or puzzled or allowed to feel that they have made fools of themselves, will give up number and seek success in some other sphere. In the course of a N.F.E.R. survey, a teacher raised the interesting point that the greatest problem in the teaching of arithmetic is that of overcoming the child's fear of the subject, occasioned not so much by his present sums, but by the fear that he will be unable to cope with the forbidding-looking ones he knows he will be taught in the future.

Valuable as these opinions are, however, they are opinions only. Most of the research bearing directly on the matter has been performed by the simple expedient of asking the subjects—students or children—what their reasons were for liking or disliking arithmetic. This technique, for obvious reasons, is satisfactory only up to a point: Wall (1958) points out that very often these emotional blocks to learning are unconsciously determined, and therefore the pupil might be quite unaware of the reasons for his attitudes. Pritchard (1935-36) shows that boys and especially girls disliked arithmetic because of a feeling of incapacity and strain when dealing with difficult items in the curriculum. Freeman (1948) bears this out almost exactly—inability to master technical difficulties he quotes as the most frequently expressed reason for not liking arithmetic. Freeman does note, however, a general improvement in attitudes over the years, and it has been suggested that this may be due to a lightening of the syllabus and improvement in teaching techniques apparent since Pritchard's work in the 1930's. Gibby (1955) found that the majority of boys and girls he studied did not find arithmetic difficult, but that those who did were more likely to dislike the subject than those who did not.

Two of the studies by Dutton are of interest here. In the earliest (1951), when his subjects were student teachers, the most frequent reasons given for disliking arithmetic were, in order: never taught the reason why; disassociated from life and social usage; pages of word problems; boring and unnecessary drill. Later (1956), when children were asked their reasons, the order became slightly changed: don't understand; too difficult and complicated; not good at it (" I don't learn easily "); boring and repetitive. The children's reasons for liking the subject emphasised, first and foremost, practical application and social needs, then followed reasons concerning the interest and fun which arithmetic had for them; only later did they mention " can work arithmetic easily . . . good at it." In a pilot study, Poffenberger and Norton (1956) interviewed a small group of 16 university students to explore previous influences upon attitudes towards arithmetic and mathematics. Some interesting hypotheses emerged. These authors, with Lovell and White (q.v.), place a lot of emphasis on parental influence which, they say, determines the child's initial attitude and affects his achievement. Where parents used themselves to have trouble with mathematics, so does the child, although this may not apply where they have favourable attitudes to their child's study generally. Poffenberger and Norton, in addition, claim that the personality of the teacher, and whether he is liked or disliked, also influences the child's attitudes.

**T.E.S.*, October 10, 1958.

This relationship between dislike of teacher and dislike of subject is worth considering a little more closely. Most people would probably imagine that there is a close connection between these two factors: the indications from research, on the contrary, generally show that this is not as clearly so as might be expected. Pritchard asked a section of his sample if their best-liked subjects were taught by the teachers they liked best, and their worst-liked subjects by those they liked least. The weight of the evidence, Pritchard says, is " emphatically against the opinion that the popularity or unpopularity of the teacher is the main influence." Pritchard was here concerned with school subjects in general, but later workers are mostly of the opinion that it is just as true when only arithmetic is concerned. Lovell and White, Freeman and Gibby are all agreed that his attitude to his teacher does not appear to have more than a small influence upon the pupil's attitude to arithmetic. On the other hand, E. E. Biggs (1951), in a study of backwardness in arithmetic, says that the pupil's attitude to the subject was of far-reaching importance, but claims that this is, to a large extent, dependent upon the relationship to the teacher. The fact that Biggs was working in a girls' grammar school, where personal relationships to the teacher are likely to be of greater importance, is a significant point in interpreting this finding.

An interesting finding in an N.F.E.R. pilot study was that amongst a group of 54 third-year junior school children, there was a highly significant correlation between liking for teacher and attitudes to arithmetic in the case of the boys in the sample. When it came to actual arithmetical performance, however, the correlation between boys' attitude to teacher and arithmetic attainment test results was not different from zero. Non-significant correlations were obtained for girls in both attitude to teacher with attitude to arithmetic, and attitude to teacher with arithmetical performance. In the case of attitudes to, and performance in, English and liking of teacher, however, no significant relationships could be found at all. These findings bear out those of the other investigators in that the attitude to the teacher does not appear to affect, at this age level, either boys' or girls' performance, although the boys tended to dislike arithmetic when they disliked their teacher and *vice versa*. Obviously, we are nowhere near the heart of the matter. It is very likely that different results would be obtained at different age levels—particularly in the difference between boys and girls. Further, liking for the teacher is closely related to the fear of failure which the child may feel—the more he likes his teacher, the more anxious he, or she, will become that he may fail and thus offend the teacher. As we shall see later, the effect this anxiety will have upon the child's work is unpredictable: in some cases it may result in improved performance, while in others it will interfere with and impair performance. It is appropriate to mention here that this sort of attitude, and particularly the teacher's own attitudes, seems to account very largely for sex differences in mathematical performance which are frequently noticed in the latter years of the secondary school.

It might be remarked that the research findings reviewed in this section do not, in the main, appear to deviate very far from what is told us by common sense and sound experience. The crucial period in developing attitudes in the child seems to lie in the infant school where his first formal acquaintance with arithmetic is made: thereafter, he should proceed at a pace commensurate with his understanding, so that he is not baffled on the one hand, or bored on the other. This sounds very simple and almost too obvious to put down upon paper, yet in how many schools does this actually happen?

The Effect of Attitudes upon Mathematical Performance

There remains the important question about the effect attitudes actually have upon the child's performance, and indirectly, upon his subsequent choice of studies or occupations involving mathematics. The latter is an important rider to the first question since it is well known that children or students can swallow their dislike of a particular subject if by so doing they can achieve some ultimately desirable end. Is it, for example, a fact, as *The Times* suggests, that the shortage of " top " scientists is to be attributed to a distaste for mathematics produced by deplorable teaching in the primary school? One recent study, reported by Lovell and White (1958), produces some evidence favouring this view. Over one hundred training college students were asked a series of standardised questions about their reasons for choosing arts or science subjects at college. Amongst the conclusions of this study, the following are relevant:

(*i*) All the students tended to study at tertiary level those subjects which they had enjoyed the most, or performed best in, during their later years at grammar school—although most students seemed to have made up their minds before they were 14 years old, which is long before they had to make a final choice.

(*ii*) Parental interests, classified as either scientific or artistic, were found to be significantly related to the choice of college career made by college students.

The following table summarises the most important findings:

Table I

REASONS FOR CHOICE OF ARTS OR SCIENCE COURSES.
102 TRAINING COLLEGE STUDENTS.

Male Students	Liked arithmetic in junior school, or was successful	Cannot remember or average	Disliked arithmetic in junior school or was unsuccessful
Non-Science . .	13	22	20
Science	31	14	2

The difference between the science and non-science groups is significant and leads the authors to draw the conclusion that the quality of arithmetic teaching in junior school should be regarded as of major importance in determining the quantity and quality of science teachers and graduates generally who are ultimately produced.

These data support the view expressed in the report by the ATCDE (1954) which states that " an attitude towards the subject (mathematics) is established several years before the children reach secondary school " and therefore " it is important that a good attitude towards, and attainment in, arithmetic should be built up in the junior school."

It is unfortunate, however, that Lovell and White did not differentiate liking and disliking on the one hand, and level of performance on the other. Thus, the general trend of their results could be expected on the basis of the success-failure criterion alone. Nevertheless, their study is important because it does emphasise the fact that success, at least, in pre-secondary school arithmetic is closely related to the choice of a science course at the tertiary stage. Other research studies have thrown further light upon this relationship between liking and performance, among which are those specifically investigating the causes of retardation in arithmetic. Thus, Wheeler (1927) states that emotional factors seem to be " most potent in producing and maintaining a condition of backwardness " and that these are difficult to correct. Drummond (1922), in discussing children specifically retarded in mathematics, claims that the difficulty and its solution are to be found in the region of the emotions. Similarly, Hoel (1954) found that lack of success in mathematics amongst intelligent children was in most cases due to emotional disturbance which could be manifested in two ways: anxiety about making mistakes and receiving reprimands, and (what probably amounts to the same thing) feelings of being under pressure from parents and teachers.

The Schonells (1956) discuss the importance of emotional attitudes in relation to backwardness: " . . . disturbing thoughts or emotional attitudes due to anxiety about himself or family matters . . . all exert a more powerful influence on attainments in arithmetic than on those in any other school subject." Evidence from factor analytic studies was cited at the commencement of this article. In particular, Hamza (1952) found that an " attitudes factor " was especially important in the case of retarded children. Finally, Burt (1949), says that the excitement, shock and general disturbance occasioned by the war upset all school subjects, but since arithmetic calls upon the most delicate functions of the mind, it became " the Belgium of the school curriculum. It suffered first; it suffered most; it suffered more conspicuously than all."

Jordan (reported in Davidson, 1950) found that the correlation between attitude and attainment was highest in mathematics, as compared with other school subjects, but even then was only of the order of 0.3. Thus, attitude, Jordan states, is a factor " helping to determine relative attainment, and attainment affects attitude, but each is the resultant of a large number of contributory causes." Davidson himself then went on to compare the relative effect of attitude upon attainment in arithmetic and English, and found that those who disliked these subjects fell below the attainments of those who liked them. This is not very surprising: the interesting point is that, in the case of English, those who disliked the subject tended to be of lower intelligence than those who liked it, and thus would be *expected* to perform at a lower level. In arithmetic, on the other hand, intelligent children no more tended to like the subject than unintelligent children tended to dislike it and therefore he comes to the conclusion that liking and being good at a subject is an evidence more in relation to arithmetic than to English.

In a small pilot study in the N.F.E.R. research programme, a third-year junior school sample was administered with a sentence completion test and the attitudes to arithmetic and English thus obtained were correlated with attainment scores in these subjects (Table II).

Table II

THE RELATIONSHIP BETWEEN ATTITUDE AND ATTAINMENT IN ARITHMETIC AND ENGLISH.

Correlations. Attitude with Performance	Boys N=23	Girls N=31	Boys and Girls N=54
English	-·14	·42*	·15
Arithmetic	·45*	·34	·41**

*=significant at 5% level.
**=significant at 1% level.

As can be seen from the above table, the only highly significant correlation is between favourable attitude to arithmetic and arithmetical performance. Such is not the case in English, where the correlation for the whole group is not different from zero—although it did approach significance with the girls. There is, however, a significant sex difference between boys' and girls' correlations in English, but this is not so in arithmetic. In fact, with a larger sample, the girls' correlation in arithmetic would also be significant—as is indicated when their results are combined with those of the boys. This sample was far too small to permit any but tentative conclusions, although these results do support those of other investigators in that the pupil's attitude to his work in arithmetic has more bearing upon his actual attainment than is the case in other subjects.

Direct evidence, then, from those studies which compare the achievements of children liking with those disliking arithmetic, and indirect evidence from studies of retardation and failure, indicate fairly definitely that poor attitudes have an adverse effect upon performance, and that this effect is relatively greater in arithmetic than in other school subjects. Although there are some grounds for believing that attitudes may be changed, negative attitudes especially seem to be lasting; at least from primary school to university level.

The Influence of Basic Personality Variables Upon Performance

We have taken a quick look at some of the secondary emotional factors. We shall now study in more detail the situation outlined at the beginning of this article where it was suggested that certain inherent personality variables such as instability and anxiety adversely affected mathematics.

We may first look at the evidence from studies of maladjustment. Feinberg (1947) summarises research relevant to maladjusted boys committed by courts to institutions. His findings agree with those of previous investigators (Chase, 1932; Armstrong and Heisler, 1945) in that a consistent pattern of scholastic achievement was found—reading and general achievement approximately average, or only slightly below the average, whereas arithmetic was very poor indeed by comparison. It was further found that intelligence had little to do with frequency of arithmetical errors, or with the arithmetic/reading ratio. Lane and Witty (1943) came to rather similar conclusions. They found that delinquents, by and large, were not noticeably inferior intellectually, but were so in educational achievement, especially in arithmetical computation. This latter was markedly improved after special coaching; more so than was the case in other subjects.

Barakat (1950), in a factor analytic study of mathematical ability, found that emotional stability, as rated by teachers of secondary grammar pupils, correlated to a small but significant extent with mathematical attainment; instability seemed to be most highly related to inaccuracy of computation. Plank (1950) studied a small group of children who were good at all school subjects except arithmetic and came to the conclusion that achievement in arithmetic in the case of these retarded subjects is more closely related to personal adjustment than to either intelligence or to school experience. She amplifies this a little by saying that maternal overprotection seems to be a most important factor in the adjustment pattern. Overprotected children may be either independent, assertive and over-critical, or overly anxious and withdrawn. Both types tend to use reading as a flight from reality, whereas arithmetic, representing reality, as it were, is deliberately avoided. Plank goes on to say that insecure children show a definite discrepancy between reasoning and computational scores, since the latter tend to be competitive and emphasise speed. Jastak (1941) also notes that the neurotic and disorganised child is usually more proficient in reading than in arithmetic.

Plank's work ties in extremely well with Levy's (1943) research into maternal overprotection. Despite wide economic and intellectual differences in his sample, Levy found constantly that in his educational record, the overprotected child achieved a high degree of success in English, vocabulary and reading, and very poor arithmetic results. The higher verbal achievement is put down to several factors, among which the most important are: (*i*) constant association with adults and " adult language "; (*ii*) consequently less time spent with a peer group which itself prohibits the use of " big words " and descriptive language; (*iii*) as there is less time spent with companions, more time is spent in solitary pursuits such as reading and study.

None of these considerations apply so readily to arithmetic, where vocabulary and reading ability are at a minimum. Thus, Levy argues, the aggressive overprotected child, having derived great satisfaction out of his high marks in English, is back to the same level in arithmetic as any other child; he would, therefore, not take kindly to a subject which did not offer him the same feeling of superiority. Secondly, mothers are not so able to help the child in arithmetic homework, especially in advanced grades. Finally, children who run errands, take paper rounds, and generally mix with other children, probably have more experience and facility with number concepts. Burt (1949) also remarks on this in his revision of the Binet scale. On the other hand, although most of the overprotected children are severely retarded in arithmetic, there are also a few who are more advanced. These latter probably themselves come from mathematical homes. The point is that no overprotected child was " average " in arithmetic and the majority were distinctly below average.

Magne* selected two groups of 'dyscalculic' children (those specifically retarded in mathematics) from a total of over 6,000 children in Swedish primary and secondary modern schools (aged 8-16), and found that the number of maladjustment symptoms was very much higher in the dyscalculic than in the normal group, especially in the case of " indolence traits " (inhibition, daydreaming, " laziness "). Of particular interest is Magne's finding that dyscalculic children of high intelligence tended to be very much more maladjusted than dyscalculic children of low intelligence. He concludes, therefore, that the arithmetical difficulties of the latter are due simply to this lack of intelligence, whereas pupils of average and high intelligence have difficulties in arithmetic only if they are maladjusted.

Instability and maladjustment are difficult and complex variables and thus it is difficult to conclude from the above studies anything further than the simple fact that children who are emotionally disturbed would be expected to show up poorly in arithmetic. Nevertheless, there is one important corollary which teachers might note. This discrepancy between arithmetic and other subjects, especially in intelligent children, is *not* to be rectified by the easy assumption that such a child is simply " lazy " in arithmetic —that he " could do it if he wanted to." Certainly, arithmetic is within the power of these children, but driving is not the way to make them realise their potentialities. Unfortunately, it is often quite difficult to distinguish the lazy from the genuinely disturbed child: it is impossible by taking into account scholastic factors only.

Anxiety is often a specific symptom of some forms of maladjustment and it has been found, generally speaking, that there, too, arithmetic is the first subject to be affected. Hanfmann (1950) points out that anxiety affects markedly both attention and concentration. Thus, immediate repetition of digits, and, to a lesser extent, arithmetic problems, are items in psychodiagnostic tests which are considered to be the most indicative of anxiety when the subject does poorly in them in comparison to other item-types.

This feature was used by Foster (1951) to investigate the relationship between arithmetic and anxiety. A group of 20 children retarded in arithmetic was matched for age and intelligence with a group of 20 controls who were not retarded. The sub-test scatter of scores of the experimental group, on the Wechsler-Bellevue Scale was like that of the typically " neurotic " pattern. Also, the vocabulary scores of this group were significantly higher than those of the controls. This study is inconclusive, but does confirm the disparity between English and arithmetic attainment noticed by Levy and others in maladjusted and presumably anxious children.

Lynn (1957) investigated the hypothesis that " good reading in relation to arithmetic might be associated with nervousness and in particular with anxiety." He found, in a sample of unselected junior school children, that there was a significant relationship with the good reading/poor arithmetic pattern. But for various reasons, and in particular as Reed and Schonfield (1958) point out, because Lynn's definition and measurement of anxiety are not unambiguous in meaning, the exact interpretation of his findings is not clear. Lynn himself seems to be interested in his results more from the point of view of reading, although he does observe that anxiety has a disorganising effect on arithmetic.

---

*Private communication to the author.

McCandless and Castaneda (1956) administered a children's anxiety scale to 359 Iowa school children and, contrary to Lynn, they found that *all* the more complicated skills—reading, arithmetic and general school performance—were apparently more vulnerable to anxiety than were the simpler mnemonic skills such as spelling. Girls were more highly anxious than boys (or at least were more prepared to *admit* their anxiety in a questionnaire) and there was generally more interference with their performance—except, and this is interesting, in arithmetic. The correlations between anxiety and arithmetic were highest in the top grades of the school (12+): —.74 (boys) and —.57 (girls), both of which are highly significant.*

On the other hand, Kerrick (1955) administered the adult form of the anxiety scale employed by McCandless, with several scholastic tests, to a group of Air Force trainees. The highest correlations with anxiety were in the case of word knowledge (—.40) and reading comprehension (—.32): the lowest was with arithmetic (—.27). These data are, of course, in direct opposition to Lynn's hypothesis. Heath (1956), using quite a different approach with a group of schizophrenic patients, in whom anxiety could be easily aroused and identified, tested verbal and arithmetical performance under stressful and non-stressful conditions. He found that performance under stress was affected much more in verbal than in number situations and he concludes that " over-learned habitual operations such as addition and subtraction may be highly resistant to disruption by even severe anxiety." However, his subjects were adults, in whom the four rules may be over-learned (they would not be in many children); and secondly Heath's choice of highly abnormal adults probably restricts the generality of his findings.

Sarason and his colleagues at Yale have reported many studies on anxiety, with special reference to their " Test Anxiety " scale. This is an attempt to measure the degree to which children are upset by tests and examinations. Some of these reports (1955, 1958, 1959) mention the relationship between test anxiety and mathematical achievement. Generally speaking, there have been no striking results except that children considered to be reading, arithmetic and behaviour problems obtained significantly higher test anxiety scores than normal children. It should be noted that there was little or no tendency for the problem children in arithmetic to be more anxious than either of the other two problem groups.

Sarason's scale has been adapted and modified for use with English children† so that three anxiety scores were available—general anxiety, test anxiety and number anxiety (i.e. the anxiety displayed when the child is concerned with numerical operations). This was administered to a group of 79 third-year children and these results were compared with arithmetic, English, and to follow up Lynn's work, the difference between English and arithmetic attainment scores. Zero correlations were obtained between arithmetical performance and all three anxiety measures. There was a small but significant correlation between the anxiety measures and English performance, a test in

---

*Because there is an *inverse* relationship, " high " correlations mean here high *negative* correlations.

†Pilot study, N.F.E.R.

which reading ability played a large part, and this is confirmed by Lynn (1955). However, the general anxiety measure did not correlate significantly with the disparity between English and arithmetic as would be expected. On the other hand, there was a highly significant correlation between the number anxiety scores and the arithmetic-English disparity. This changes the focus of Lynn's original contention that general anxiety inhibits arithmetical performance and promotes reading ability: the present data indicate that number anxious children, particularly girls it was found, expend more effort in English in the attempt to find some security in their studies at school. This is a reaction which, as we have seen, Levy discovered in over-protected children.

The apparent inconsistencies in all these results may be explained to a certain extent if we recall that the relationship between anxiety and performance is curvilinear. When anxiety is very low, that is, performance tends to be poor simply because there is a level when the subject " couldn't care less " and obviously in this condition he will not perform at his optimum level of efficiency. On the other hand, under extreme conditions of anxiety, performance very often becomes utterly disorganised. Sullivan (1951) found that at crucial examinations, when anxiety should be experienced and when motivation was high, bright children especially did very much better in arithmetic than when anxiety was relatively less during the school year. So much depends upon the subject, the task, and the kind of anxiety experienced. For a more detailed examination of the relationship between anxiety and performance, the reader is referred to Eysenck (1958).

Number Anxiety

We have now considered both primary and secondary emotional factors which may be operative in mathematical performance. Generally speaking, it was suggested that the secondary factors are the result of external conditions such as teaching technique and are thus theoretically within reach of correction. Primary factors, on the other hand, are deeply rooted in the personality of the child and nothing short of psychotherapy can do very much about them. It is nonetheless important that teachers should be aware of the difficulties which anxious and disturbed children are likely to find in arithmetic so that, in the early stages especially, they are helped over the obstacles without being castigated for laziness or stupidity. Very often, they are neither lazy nor stupid—but may well appear to become so in the absence of understanding and guidance.

There is, however, another emotional factor in mathematics which appears to possess some primary as well as some secondary characteristics in that it is pathological in intensity and at the same time specific to mathematics. It seems to go very much deeper into the personality pattern than do simple feelings of distaste, and it also appears to be much more closely related to basic personality variables, anxiety in particular. This is what Dreger and Aiken (1957) aptly call " number anxiety."

Now it is evident that this is not the same factor as causes the poor mathematical performance of the anxious or unstable person, nor that which intervenes in the person who has poor attitudes to the subject. The number anxious child, or adult, is by definition deeply disturbed when he is dealing with number and arithmetical operations: the child who only dislikes arithmetic, however thoroughly, may not feel particularly anxious or insecure in his performance.

There is only one piece of research, apart from the N.F.E.R. studies already mentioned, that has made an investigation into this specific problem, and this is reported by Dreger and Aiken. They hypothesised that number anxiety was a factor distinct from general anxiety, that it is uncorrelated with I.Q., and that number anxious persons would perform at a lower level of mathematical performance than non-number anxious persons of the same intelligence. The Taylor anxiety scale, with three additional number items, was administered to a student population and from the results, subgroups of students were selected at random from each of the four extremes of the scores. Each subgroup was tested with the Wechsler intelligence scale and with each subtest, a galvanic skin response reading was taken (the so-called " lie-detector," which is extremely sensitive to autonomic anxiety reactions). Only on the arithmetic subtest were there any significant G.S.R. readings, and these clearly differentiated the number anxious from the non-number anxious subjects. Inter-correlations between number test items were high, those between number and general anxiety items were low, thus suggesting that number anxiety is a separate factor in its own right. There was no clear relationship between number anxiety and intelligence, but there was a significant negative correlation with mathematics results in university examinations. The authors, on the first point, conclude that low intelligence *per se* is not a factor causing number anxiety. This important study, however, can only be regarded as suggestive, especially in view of the very small number (3) of test items, and because the subjects were undergraduates, not children.

Summary and Conclusions

It was found convenient to classify the research findings bearing upon the emotional factors affecting mathematical performance under three headings: primary, which are part of the basic personality structure; secondary, which are specific attitudes of liking or disliking, but which may, of course, have their foundations in primary factors; and finally, there was a third kind of factor which seems to share certain characteristics of both primary and secondary factors.

Many studies report a relationship between maladjustment and poor mathematical performance. Although this is also true of some other school subjects, this relationship seems to be much closer in the case of arithmetic. It is not suggested for one moment that some people may have an innate fear of mathematics, nor even that a specific mathematical " ability " is lacking in some. Rather the evidence suggests that certain *general* personality traits, such as anxiety, instability and so forth, inhibit the person possessing them from fully realising his cognitive potentialities—at least, when mathematics and number are taught as they have been hitherto. It may well be that the maladjusted child would take quite happily to mathematics if different methods of introducing the subject to him were used. This is a point which needs to be demonstrated by future research: the fact that intelligent dyscalculic children *can* be made mathematically proficient when they receive appropriate remedial teaching indicates that this hypothesis might be true. Many so-called " lazy " children are those who, because of their temperamental characteristics, have found an initial block to their number learning which has been heightened by accusations of laziness (because of their otherwise obvious intelligence) and hence by unimaginative driving on the part of their teachers—which, of course, disturbs them even more. The vicious circle thus set up may often result in the phenomenon known as number anxiety, which is anxiety specifically directed towards arithmetic and which is more devastating in its effects upon performance than simple disliking, although naturally distaste is also involved. However, research into this particular topic is too meagre so far to allow anything but inspired guess-work.

The research results bearing upon straight distaste, as distinct from definite fear, are rather more fruitful. Attitudes appear to arise as early as the infant school stage, but may be deepened or changed probably at any point during the child's school career. The most common reasons for negative attitudes are based upon a lack of a proper understanding of the subject, which is in many cases due to mechanical teaching; and upon plain boredom, a product of uninteresting teaching. Parental and teacher influences are also of importance but the influence of the teacher is more likely to depend upon his own attitude to the subject he, and especially she, teaches (how many teachers themselves really *like* arithmetic?) rather than upon the child's personal liking for him or her. Not unexpectedly, children with poor attitudes to arithmetic do very much worse than those who like the subject; moreover, the influence of attitudes upon performance is greater in arithmetic than in English or other subjects. The child who dislikes English may yet achieve high marks in that subject—this is not so likely to occur in arithmetic.

The implications of all this for the teacher are obvious, and need no further elaboration. There is no intrinsic reason why mathematics should be distasteful. As it evidently is so for many people, the fault must lie in the classroom. We do not these days hear anything about a " Fear of English," or a " Fear of Latin." How long will it be before mathematics is put in its rightful place alongside these other subjects so that it will become just as meaningless to speak of the " Fear of Number "?

References

Armstrong, E. and Heisler, F., " A note on the attainment of delinquent boys." *School and Society*, 61, 29-31, 1945.
A.T.C.D.E., *Arithmetic in Primary Schools*. London: Longmans, Green & Co., 1954.
Barakat, M. K., " Factors underlying the mathematical abilities of grammar school pupils." Summarised in *Brit. J. Educ. Psych.*, 21, 239-240, 1951.
Bartemeier, L., " A counting compulsion." *Int. J. Psychoanalysis*, 22, 301-309, 1941.
Biggs, E. E., *An Investigation into the Causes of Backwardness in School Mathematics*. M.A. thesis, University of London, 1951.
Biggs, J. B., " The teaching of mathematics—I: The development of number concepts in children." *Educ. Res.*, 1, 17-34, February, 1959.
Burt, C., *Mental and Scholastic Tests*. London: Staples Press, 1949.
Chase, V., " Educational achievement of delinquent boys." *J. Juvenile Res.*, 16, 189-192, 1932.
Churchill, E. M., " The number concepts of the young child." *Researches and Studies*, University of Leeds Institute of Education, No. 17, 34-49; No. 18, 28-46, 1958.
Davidson, R., *The Relation between Attitudes and Attainment in various aspects of Arithmetic and English as shown by Senior Pupils in the Primary School*. B.Ed. thesis, University of Aberdeen, 1952.
Dreger, R. R. and Aiken, L., " The identification of number anxiety in a college population." *J. Educ. Psych.*, 48, 344-351, 1957.
Drummond, M., *Psychology and Teaching of Number*. New York: World Book Co., 1922.
Dutton, W. H., " Attitudes of junior high school pupils towards arithmetic." *The School Rev.*, 18-22, January, 1956.
Dutton, W. H., " Attitudes of prospective teachers towards arithmetic." *Elem. Schl. J.*, 52, 84-90, 1951.
Dutton, W. H., " Measuring attitudes towards arithmetic." *Elem. Schl. J.*, 52, 24-31, 1954.
Eysenck, H. J., *The Dynamics of Anxiety and Hysteria*. London: Routledge and Kegan Paul, 1957.
Feinberg, H., " Achievement of a group of socially maladjusted boys as revealed by the Standford Achievement Test." *J. Soc. Psych.*, 26, 203-212, 1947.

Foster, P. M., *An Examination undertaken with School Children of the Hypothesis that Anxiety encroaches ont he Attention in such a way as to impair Performance in Arithmetic.* B.Ed. thesis, University of Belfast, 1951.

Freeman, W. A., *An Enquiry into the Attitudes of Secondary School Boys and Girls Towards Arithmetic.* M.A. thesis, University of Birmingham, 1948.

Gibby, W. A., *To Investigate the Extent and Range of Interest in Arithmetic in Selected Groups of School Children.* M.A. thesis, University of London, 1955.

Hamza, M., " Retardation in mathematics amongst grammar school pupils." *Brit. J. Educ. Psych.*, 22, 189-195, 1952.

Hanfmann, E., " Psychological approaches to the study of anxiety," in *Anxiety*, Eds. Hoch and Zubin. New York: Grune and Stratton, 1950.

Heath, D., " Individual anxiety thresholds and their effect on intellectual performance." *J. Abnorm. and Soc. Psych.*, 52, 403-408, 1956.

Hoel, L., " An experiment in clinical procedures for arithmetic." In *Emerging Practices in Mathematics Education*, 22nd Year Book of the National Council of Teachers of Mathematics, Washington, D.C., 1954.

Jastak, J., *Wide Range Achievement Tests.* Wilmington: Story, 1941.

Jordan, D. See Davidson, 1952.

Kerrick, J., " Some correlates of the Taylor Manifest Anxiety Scale." *J. Abnorm. Soc. Psych.*, 50, 75-77, 1955.

Jeffreys, H., " The unconscious significance of numbers." *Int. J. Psychoan.*, 27, 217-223, 1936.

Lane, H. A. and Witty, P.A., " The educational attainment of delinquent boys." *J. Educ. Psych.*, 25, 695-702, 1934.

Levy, D., *Maternal Overprotection.* New York: Columbia University Press, 1943.

Lovell, K. and White, G., " Some influences affecting choice of subjects in school and training college." *Brit. J. Educ. Psych.*, 28, 15-24, 1958.

Lynn, R., " Temperamental characteristics related to disparity of attainment in reading and arithmetic." *Brit. J. Educ. Psych.*, 27, 62-67, 1957.

Lynn, R., " Disparity of attainment in reading and arithmetic." (Reply to Reed and Schonfield). *Brit. J. Educ. Psych.*, 28, 271-276, 1958.

Lynn, R., " Personality factors in reading achievement." *Proc. R. Soc. Med.*, 48, 996-998, 1955.

Mace, C. A., *The Psychology of Study.* London: Methuen and Co., Ltd., 1932.

McCandless, B. and Castaneda, A., "Anxiety in children, school achievement and intelligence". *Child Development*, 27, 279-382, 1956.

Ministry of Education, *The Teaching of Mathematics in Secondary Schools.* Pamphlet No. 39, H.M.S.O., 1958.

Plank, E., " Observations on attitudes of young children towards mathematics." *Math. Techr.*, 43, 252-263, 1950.

Plank, E. and Plank, R., " Emotional components in arithmetical learning as seen through autobiographies." In *The Psychoanalytic Study of the Child*, 9, 274-293, 1953.

Poffenberger, T. and Norton, D., " Factors determining attitudes towards arithmetic and mathematics." *Arith. Tchr.*, 3, 113-116, 1956.

Pritchard, R., " The relative popularity of secondary school subjects at various ages." *Brit. J. Educ. Psych.*, 5, 157-179, 1935; 229-241, 1936.

Reed, G. and Schonfield, D., " Disparity of attainment in reading and arithmetic." (Criticism of Lynn's hypothesis.) *Brit. J. Educ. Psych.*, 28, 271-276, 1958.

Russell, B., *Mysticism and Logic.* Pelican Books, No. A270, 1953.

Sarason, S. B., *Measurement and Correlates of Test Anxiety in Children.* Report on Project M-712, United States Public Health Service, June 1, 1954—May 31, 1955. (Mimeographed.)

Sarason, S. B., *et alia.*, " A test anxiety scale for children." *Child Development*, 29, 105-113, 1958.

Sarnoff, I., Sarason, S. B., Lighthall, F. and Davidson, K., " Test anxiety and the ' Eleven-Plus ' examinations." *Brit. J. Educ. Psych.*, 29, 9-16, 1959.

Schonell, F. J. and Schonell, F. E., *Diagnosis and Remedial Teaching in Arithmetic.* London: Oliver and Boyd, 1957.

Stern, C., *Children Discover Arithmetic.* New York: Harper and Brothers, 1949.

Sutherland, J., " Pupils' arithmetical ability in secondary schools." *Brit. J. Educ. Psych.*, 21, 3-8, 1951.

Wall, W. D., " The educational interests of a group of young industrial workers." *Brit. J. Educ. Psych.*, 15, 127-132, 1945.

Wall, W. D., " The wish to learn—research into motivation." *Educ. Res.*, 1, 24-38, 1958.

Wegrocki, H., " A case of number phobia." *Int. J. Psa.*, 19, 97-99, 1938.

Wheeler, E. C., *Backwardness in Arithmetic amongst London Elementary School Children.* M.A. thesis, University of London, 1927.

Williams, A., " Number readiness." *Educ. Rev.*, University of Birmingham Institute of Education, 11, 31-46, 1958.

Wrigley, J., " The factorial nature of ability in elementary mathematics." *Brit. J. Educ. Psych.*, 28, 61-78, 1958.

RESEARCH ON TEACHING ABILITY

by K. M. Evans

(Lecturer and Research Officer, Faculty of Education, University College, Cardiff)

I. General introduction

THE most important people in any educational system are the teachers in the classrooms. It is true that children do not receive the whole of their education from teachers in school, but teachers do come into contact with and help to educate far more children than any other group of people, and their influence on the rising generation can be very great and must not be underestimated.

Most people have opinions about schools and teachers, often based on memories of their own schooldays or the information brought home from school by their children. Any conversation on the subject brings out the differences and the similarities in the views of various groups, and shows up the implicit assumptions on which judgments rest. Most parents want their children to have at least as good a start in life as they themselves had, and in judging teachers and schools they take into account examination results, discipline, the character and personality development of pupils and skill in handling backward or problem children. Less often the needs of the teacher as a person are recognised, and when this is done it is usually because their satisfaction or frustration is likely to affect the pupils.

It is with the teacher as a person in relation to the children and their needs that training colleges are mainly concerned, and one of their major problems is the satisfactory assessment of the teacher's success in his or her work. Explicit, not implicit, assumptions are needed here, and there should be constant appraisal of methods of selecting, training and assessing students.

In view of the importance and interest of the subject, it is not surprising to find that a very large number of investigations have been carried out with a view to determining what kinds of people make the best teachers, and what are the effects of various personality and other characteristics on the efficiency of the teacher. There have been a number of reviews of this research published, the fullest of which covers over a thousand studies. These reviews show that since 1931 there has been an interesting shift of emphasis from a preoccupation with the importance of qualities such as intelligence, scholarship and professional knowledge, to more descriptive studies of the teacher's personality and relationships between pupils and teachers.

This research is of importance in that it provides information which is of use in the selection of candidates for training as teachers and in the prediction of the probable level of success of potential teachers. It can also indicate the probable outcomes of teaching methods and thus enable courses of training to be suited both to the capacities of the students and to the needs of their future pupils.

II. The assessment of teaching

The problem of how to assess teaching ability is fundamental to any research on the subject, and the value of any piece of research will depend on the soundness of the criterion used as well as upon the design and carefulness of the study. At the outset a distinction must be made between teaching ability and teaching efficiency or competence. Too often they have been confused and the terms used as if they are interchangeable. This is not the case. The possession of high teaching ability does not ensure that the individual concerned will become a highly efficient teacher, or even that he will become a teacher at all. It only means that, given the desire and the opportunity to do so, he has the capacity to become a good teacher. It is at least questionable whether research on teaching ability is possible. Certainly, existing research really deals with teaching efficiency, though it is frequently described as teaching ability.

The importance of the distinction is different in different contexts. Students who are being selected for training as teachers usually have little or no teaching experience, and in these circumstances it is impossible to judge their teaching efficiency. Consequently, selectors attempt to assess their teaching ability and recommend for training those who appear to have the necessary characteristics for success. Studies of the results of these procedures often produce evidence of their reliability by considering the extent to which different selectors agree, but, since no evidence is available of the subsequent efficiency as teachers of those rejected for training, sound information on their validity is lacking. There are some studies which deal with the efficiency of those selected for training in relation to the criteria by which they were selected, and these are often cited as validation studies. They are unsatisfactory because they deal only with a selected population.

When appointing trained or experienced teachers to posts there is some record of their efficiency either as student teachers or in previous posts. In this case, their known level of efficiency is taken into account rather than their potential based on an assessment of their teaching ability.

In all this there is the implicit assumption that it is possible to assess teaching efficiency with some degree of accuracy. Certainly, assessments of teaching are made fairly frequently. They are made by people concerned with the training of teachers and by inspectors when they watch teachers in the classroom. They are made by head teachers who form opinions about the efficiency of members of their staffs, and, less formally, by pupils and their parents. Some of these assessments are recorded as examination marks or in testimonials or professional records, and they are usually treated as having some meaning, though occasionally one becomes aware of doubts about them. How do we, in fact, assess teaching efficiency and are our methods satisfactory?

One way of assessing teaching efficiency is by considering the results of teaching. Teachers must teach something. How much of this have the children learnt? This can be discovered fairly easily by testing their knowledge before and after a course of lessons. The increase could be considered as a measure of the efficiency of the teaching. There are a number of studies which have made use of this measure. Undoubtedly, the teacher who adds little or nothing to the children's store of knowledge cannot be considered as being particularly efficient. On the other hand, mere cramming with facts is not all that is expected of a teacher. Everyone who has had experience of coaching children for examinations knows that it is possible for them to know temporarily a great many things which are forgotten with surprising rapidity once the examination is over. This raises the question of whether immediate or retained knowledge should be used as the criterion of teaching efficiency.

Again, the amount of knowledge gained depends not only upon the efficiency of the teacher but also upon qualities of the children. One obvious quality affecting learning is intelligence. Where the efficiency of different teachers is being compared it is essential to take into account the intelligence of the children they are teaching. The more intelligent the class the more they should learn. It is probably safe to say that a highly intelligent class will learn something from even the most inefficient teacher, whereas a dull class may learn pitiably little from even the best of teachers. If the amount of knowledge gained is to be used as a criterion of the efficiency of teaching it is essential to see that the classes taught are at least comparable in intelligence.

Other factors besides intelligence affect the children's capacity to learn. Among these one of the more important is their home background. The child who comes from a home where there are books and pictures, where he hears good conversation, whose parents are interested in his education and prepared to help him by talking to him and taking him to places of interest, has a much better foundation to build on than the child who lacks these advantages. New knowledge can be linked to old and is better remembered in this way than when it is seen as isolated information. But this result is in no sense due to the efficiency of the teacher. Amount of knowledge gained must be regarded as a dubious criterion of teaching efficiency, though it is undoubtedly one of the more measurable criteria.

Another objection to the use of increase in knowledge as a criterion of teaching efficiency is that it leaves out too much. Children gain much more than factual information from their teachers. A teacher can affect a child's whole attitude to life and can help to build up or to shatter its ideals. Heroes and cowards can be made in the classroom. But changes of this kind cannot be gauged and used to measure the efficiency of the teacher. They can, at best, be sensed by observers.

Probably the opinion of competent observers is the best criterion of teaching efficiency. Certainly, it is the one most commonly used. In arriving at a rating of any teacher the observer can take into account not only the quality of the instruction being given but also more intangible things, such as the attitudes of the teacher and the children and the whole atmosphere of the lesson. A classroom with teacher and pupils is, after all, a small society, and the way in which it functions should be considered when estimating the efficiency of the teacher. The children may be working hard and learning a great many facts, but they may be doing so either because they are afraid of a sarcastic teacher or because the teacher is making the lesson so interesting that they want to learn. This could only be distinguished by an observer on the spot, not by the results of an achievement test.

There are, however, many arguments against the use of the opinion of an observer as a criterion of teaching efficiency. There is the obvious argument that observers do not always agree among themselves. A lesson which seems very good to one observer may seem very poor or only mediocre to another. This may be because their actual standards differ, or because, in making their assessments, they are not taking into account the same qualities or are weighting them differently. It is possible to ensure that the same qualities are being considered by drawing up a list of those that are thought to be important. This can, if desired, be used as a rating scale, each quality being rated separately and a final rating derived from the results. If this is done it is usually advisable to define the qualities in as precise terms as possible. This helps to make certain that the observers are all using the terms in the same sense. The relative weights to be given to the various qualities should also be decided upon if the ratings of different observers are to be comparable.

Other difficulties which may arise in the use of this criterion are the result of the interaction of the personalities of the teacher and the observer. A teacher may easily be ' put off ' by the presence of one observer but may be encouraged and stimulated by another. Then, too, teachers may vary in their performance from day to day and from class to class. A teacher who is very good with older children may be very poor with younger ones, or vice versa. The subject matter may exert an influence. A teacher who teaches one subject well may teach another much less well. The mood of the children varies, too. On some days they are eager to learn and easy to teach; on other days they are dull and listless or bored and unwilling to learn. This is not always due to the teacher in front of them at the moment. It may be the result of something which happened in an earlier lesson, it may be due to the time of day or the stage of the term, or even to the weather. Every teacher knows that there are certain kinds of weather which upset the children and make them difficult to teach, and that lessons go better at certain times of day than at others.

Again, a teacher who is unsuccessful in one school may be successful in another. The personality of the head teacher, conditions in the school, opportunities for specialisation and other considerations of this sort have been known to affect a teacher's performance.

Much more could be written, but probably enough has been said to show the kind of difficulties which are likely to be encountered in assessing teaching efficiency. Yet here we are on far more solid ground than is the case when we are attempting to assess teaching ability. With teaching efficiency we can study the performance of the teacher, either directly or by report, but teaching ability is a potential, and as such is not directly observable. Nevertheless, we must find some way of assessing teaching ability if our selection of students for training as teachers is not to be made purely by chance.

III. Teaching ability in relation to specific qualities

The obvious way to go about studying teaching ability is to consider various qualities which it is reasonable to expect to find in good teachers and to attempt to relate them to the level of observed efficiency. The results would provide some guidance as to the qualities to be looked for in selecting student teachers.

Many people have drawn up lists of the qualities which should be considered in judging teachers. Some of these appear to be based on nothing except the personal opinions of their authors. Others represent a consensus of opinion as to qualities desirable in teachers, or are the result of the mathematical analysis of correlations between assessments of teaching efficiency and other measures. A wide variety of qualities are mentioned in these lists, but the main ones can be classified as physical characteristics, intelligence, scholarship, professional information, attitudes and interests, and personality. A more detailed discussion of the importance of each of these follows.

1. *Physical characteristics*

When lists of qualities desirable in teachers are studied, it is noticeable that while both pupils and administrators consider that the personal appearance of the teacher is important, health is generally mentioned only by administrators. In the main, health and personal appearance come low on these lists, but there are variations according to the age of the children taught and the type of school. Early correlational studies show little relationship between physical characteristics and teaching efficiency. The personal attractiveness of their teachers possibly matters more to girls than to boys and to the middle age groups in a school more than to younger or older ones. Administrators are bound to be somewhat concerned over the health of their teachers, since a teacher who breaks down or is continually absent through illness is a liability, but given the minimum level of health necessary to sustain the work, it is probably fair to say that anything over and above this minimum adds little to the effectiveness of the teacher.

Whether speech is a physical characteristic or not is open to question, but since quality of voice is, voice and speech will both be considered here. There are very few investigations of the importance of the voice of the teacher, and such as there are suggest that the correlation between voice and teaching efficiency is low, except where the two sets of ratings are made by the same people. This may be evidence of halo in the ratings, but it also suggests that voice was one of the things being taken into account when teaching was assessed.

This is justifiable since so much classroom instruction is oral. An obvious speech defect may be an insuperable obstacle to a teacher, and candidates with such defects should be advised to adopt a career where speech is less important. A strong local accent is not always a handicap. If the teacher is going to teach in his or her own area it may even be an asset, since the children will understand it and the teacher will be accepted as part of the community. An accent strange to the area can mark a teacher off as a 'foreigner', and may result in social isolation.

A poor command of language, on the other hand, may be a real handicap. A teacher who is at a loss for words may have difficulty in conveying his meaning to a class, and his descriptions and story-telling are likely to be impoverished by his lack of words. This is a particularly serious fault in a teacher of young children, who are themselves building up their vocabularies. A poor vocabulary or lack of fluency in the teacher is very likely to have an adverse effect on the speech of the children.

While the results of studies suggest that physical qualities do not make a large contribution to the efficiency of the teacher, it is probably true to say that gross defects of physique, health or speech are undesirable in a teacher who has to deal with large numbers of children.

2. *Intelligence*

There appears to be a widespread and long-felt conviction that it is important that teachers should be people of high intelligence. Intelligence appears high on the lists of qualities desirable in teachers and in the period 1937-39 it was the fourth most frequently studied quality. In 1940-43 it came second.

Results of statistical studies, on the other hand, leave the question of the importance of intelligence in some doubt. In many studies correlations between the results of intelligence tests and assessments of teaching efficiency are small, but there are other investigators who produce evidence that intelligence is more highly correlated with teaching efficiency. Bliss (1922) showed that " the chances are twice as good to have superior or excellent teachers if individuals of superior ability are employed, as when individuals of average ability are employed, and almost four-and-a-half times as good as when individuals of low rank mentally are hired." More recent studies have suggested that the general knowledge and mental ability of the teacher are among the most important factors conditioning the amount the pupils learn.

The present position seems to be that there is a minimum level of intelligence necessary for success in teaching, but that, given that minimum, there is little relationship between teaching ability and actual level of intelligence. At the same time, one writer suggests that intelligence and the willingness to use it in the education of children is the most important factor in teaching ability. This, however, introduces the idea of motivation.

There is evidence that success in training college written examinations is related to intelligence, and intelligence plays a part in the preparation and presentation of teaching material. Whether the same level of intelligence is needed in teachers of younger and older children is sometimes questioned, but since they all have to pass comparable examinations a distinction is probably impracticable even if it were desirable.

Considering how easy it is to apply an intelligence test it is surprising how few colleges make any real attempt to assess the intelligence level of candidates for admission. It is true that scholastic attainment does give some indication of intelligence level, but so many other factors are involved as well as intelligence that the two cannot be considered synonymous. No-one would claim complete accuracy for the results of intelligence tests, but as additional information about the candidates they might well be more widely used than they are in selecting students for training.

3. *Scholarship*

How much should a teacher know? Stories of brilliant scholars who are bad teachers and of plodders who are excellent teachers are only too common, but fortunately there is more objective evidence about the connection between scholarship and teaching success. In a number of studies the academic marks obtained by students during their college courses are correlated with their marks in practical teaching. The coefficients range from about 0.2 to about 0.6. They are not high, but they are positive. In the main, the conclusion to be drawn from this is that usually there is a slight tendency for students in training who are strong academically to be better teachers than those who are weak academically. Similar results have been obtained when college academic records of experienced teachers have been compared with their ratings as teachers in the field.

There is plenty of evidence in research reports in support of four statements.

1. Poor teachers tend to be lacking in scholarship.
2. Students who are lacking in scholarship tend to make poor teachers.

3. Well-informed students tend to make good teachers.
4. Good teachers tend to be well-informed.

The importance attached by pupils to their teachers' scholarship varies according to the age and mental ability of the pupils. It has been found that older and abler pupils are more impressed by scholarship than are younger and less able pupils. Possibly the latter groups assume that their teachers are knowledgeable people while the former have become more critical. The gap in knowledge between teacher and pupils is, in any case, much wider for younger pupils than it is for older ones, and for this reason if no other, high scholarship is less necessary in teachers of younger children than in teachers of adolescents.

What, then, should be the policy of those who have to select students for training as teachers? It is obvious that they must take into account the academic record of candidates since it is necessary that a teacher should have an adequate knowledge of subject matter. Some of this knowledge will be acquired during training, but for this to be done successfully the student must have a sound foundation of basic information to begin with. This should have been acquired before entry to college. The course of training is short, and during it, not only academic but also professional subjects must be taught. It is not by any means uncommon to listen to a lesson being given by a student whose teaching technique is adequate but whose knowledge of subject matter is meagre. It is possible, of course, that experience will bring with it a better grasp of subject matter. At the same time the standard of entry to training colleges is not too high, and a candidate who has not reached this before entry is likely to have difficulty with the academic side of the training course and is not likely to make a good teacher. Research findings make it clear that there is a strong case for excluding academically weak candidates from training courses.

4. *Professional information*

There are a number of ways in which a teacher's professional information and knowledge of teaching procedures can be tested. This is done during training by examinations in the general theory of education and in methods of teaching particular subjects. When the results of these examinations are correlated with training college assessments of practical teaching, the resulting coefficients are usually small but positive. At that stage, knowledge of methods of teaching is not closely parallelled by the ability to put them into practice in the classroom. It must be remembered, though, that the students have had much more experience of answering examination questions than of practical teaching. As their teaching skill improves, it may be that the connection will become closer. Certainly, it is unlikely that a teacher with no background of theoretical knowledge, however this may have been acquired, will be a very enlightened practitioner.

In addition to examinations of the type mentioned, there are in existence a number of more or less standardised tests of professional information. A common type describes various situations which may arise in school and asks for decisions on the best ways of dealing with them. Very often several possible ways are suggested, and the 'rightness' of an answer depends on the degree to which it agrees with the opinions of experts.

Some investigators claim that these tests really do discriminate between good and bad teachers, but it is more common to find that the tests in this form are too easy for the subjects, who all tend to score highly. Apparently it is not at all difficult to see what is the best course of action, but there is no guarantee that the situation, if it arises, will be dealt with in this way. It has been suggested, probably with justification, that these tests are more closely connected with intelligence than with teaching ability. This criticism might apply to all written tests of professional information, including traditional type examinations.

5. *Attitude towards and interest in teaching*

Several investigations of attitude towards teaching as a career have been made. The earlier ones were concerned with the reasons given by school children and students for the choice of teaching as a career. A later study investigated the influence on attitude towards teaching as a career of such factors as the school and the pupil's attitude towards it, intelligence, academic success, home circumstances and types of interest. Of these, the pupil's attitude to school seemed to be the most important. These studies were not of a type in which attitude could be linked with teaching efficiency as the subjects were, for the most part, still at school or in the early days of their college course.

The idea that a teacher's attitude to his work might be connected with his efficiency is not new. In 1935 Birkinshaw used satisfaction with the work as a criterion of success in it, on the grounds that it implies that the happy teacher has " both the affection and regard of many of those with whom and for whom she works." It has also been noted that maladjusted teachers tend to be critical of the system, people and conditions of work, and that while teaching procedures and techniques can be modified, feelings and attitudes are a deeper core of the personality and cannot be changed easily.

Where attitude scales have been used, it has generally been found that correlations between attitude and teaching efficiency have been low. There have also been examples of significant negative correlations between attitude scores and intelligence test results. The reasons for this may have been social and economic, in that the less intelligent subjects saw teaching as a safe career and were disposed to prefer it to something more adventurous and less safe. It is also possible that the less intelligent students are less critical of teaching conditions than more intelligent ones, and therefore tend to show a more favourable attitude.

It has been claimed that some American attitude scales differentiate well between different types of teacher, but they do not appear to do so when used in Britain. British norms for these scales seem to differ considerably from published American ones, and this is probably due to cultural differences and different ideas about children's education held in the two countries. Consequently, American attitude scales are of little use in this country.

In contradistinction to studies of attitude, there are very few studies of interest in teaching. That by Birkinshaw has already been mentioned. A test of interest in different aspects of the work of a teacher was constructed by Evans (1952a) and administered to several groups of students in training. The work was analysed under four headings as connected with the intellectual and social training of children, the organisation of the school, and relationships with colleagues. The test was presented in the form of a narrative account of a teacher's day, and the subjects were asked to assess their interest in a number of the activities mentioned. The results showed little relation between success in practical teaching and interest in these aspects of the work. In one study evidence was found that interest scores were negatively correlated with intelligence.

Interest in teaching as measured by these tests does not appear to be a good predictor of success as a teacher, nevertheless it would be folly to ignore the candidate's desire to teach when selecting students. Probably the best course is to make sure that interest is based upon some real knowledge of what the work involves, and is not merely an uncritical approach or an unrecognised desire to remain at school and avoid the strains and stresses met in the world by adults.

6. *Interests*

Belief that it is important for teachers to have plenty of interests outside their work is widespread. These interests are considered to be links which bind teachers to their pupils and to the community, and bring life and vigour into their teaching. Some writers have produced evidence that children like their teachers to have such interests, but there is also evidence that pupils consider other qualities, such as patience, friendliness, fairness, and a sense of humour, to be more important.

There are plenty of studies of the mental adjustment of teachers in relation to their interests. Teachers with varied interests, particularly of an active, outdoor or social kind, have been found to be well-adjusted, while those with few interests or interests closely allied to their work tend to be maladjusted. This might be equally true of any other work group, and it is doubtful whether it applies to teachers more than to others. Indeed, there is no real evidence that teachers have characteristic types of interests which mark them off from other occupational groups. Nor is there much to suggest that type or range of interest is connected directly with teaching efficiency.

At the same time there is evidence that personal contentment is increased by the pursuit of a hobby. The contented person, whether a teacher or not, is likely to have more emotional energy free, and therefore to be more efficient than one who is less contented. This, rather than any connection between the hobby and teaching efficiency, is a sufficient reason for encouraging students to develop hobbies outside their work, and for taking these into account when selecting students for training.

7. *Personality*

Personality is an ambiguous term and as such lends itself to many interpretations. There exists a comparable number of studies of teaching efficiency in relation to the personality of the teacher. Large numbers of personality qualities figure in the lists of teacher characteristics already mentioned, and studies dealing with them, both severally and in combination, have been carried out. Among the most frequently named are such qualities as leadership, initiative, enthusiasm, social qualities, resourcefulness, sympathy, fairness, tact, integrity and cheerfulness.

A few of these studies deal with personality qualities as measured by various tests and questionnaires. The majority, however, use ratings made by supervisors, inspectors, head teachers, colleagues, pupils or the teachers themselves. In many cases little or no correlation is found between personality ratings and teaching efficiency. Exceptions occur, usually when the two sets of ratings are made by the same people, and one cannot but suspect that this is really a halo effect, and that the ratings are not independent.

One thing which emerges very clearly from this work is that there does not appear to be any ' best ' type of teaching personality. Even poor mental health is not always a handicap. Emotionally stable students are more likely to be good teachers than less stable ones, but one study found about 25 per cent. of the former to be inferior teachers while about 52 per cent. of the latter were good teachers.

Some qualification is probably needed here as to the degree of instability. Gross disturbance is likely to have a detrimental effect on success in any occupation, but a very large number of people carry on successfully in spite of mild emotional disturbance. Indeed, it may be questioned whether anyone gets through life without encountering some emotional difficulties. The important thing is how we cope with them, and training colleges might consider that aiding their students to meet difficulties and make a successful admustment is part of their work. At the same time, children should not be exposed to the risks to their own development resulting from daily contact with a severely maladjusted teacher.

It is probably here, in the relationship between pupils and teachers, that the importance of the teacher's personality really lies. Studies of teachers as instructors show little relationship between efficiency and personality, but there is abundant evidence that the characteristics of the teacher have a marked effect on the characters and behaviour of the children. For this reason it is necessary, when selecting teachers, to consider, not only the more easily assessable qualities already discussed, but also those less tangible qualities of personality which may, in the long run, be equally important determiners of their success with their pupils.

IV. Studies making use of factor analysis

As a result of the application of factor analysis to sets of correlations between teaching efficiency and some of the qualities considered above, evidence has been produced of a number of qualities which contribute to success in teaching. As is always the case with factor analysis, the factors found depend on the assumptions made and the tests used, and the only real information produced is concerned with the relative importance of the various factors. It is, therefore, not surprising to find that no two investigators end up with the same list of factors, or that the number of such factors varies. At least two studies mention factors which, from their descriptions, appear to be properties of the rating scales used rather than qualities of the teachers or the teaching situations.

Examples of the types of factors found are:

General knowledge and mental ability.
Intelligence and the willingness to use it in the education of children.
The ability to wear each student's ' sensorial and emotional ' shoes.
The ability to appear live and interesting to children.
A ' pupil-response ' factor, positively loaded for the abilities and qualities of the teacher which dispose the pupil to respond.
A factor concerned with the teacher's ability to deal with classroom situations and to make use of his past teaching experience.

Professional maturity.
Personal, emotional and social adjustment.
Personal appearance and attractiveness of the teacher.
A speech factor.
A eulogizing attitude towards the teaching profession.

The reader may well ask what all this amounts to. Have the factor-analysts really decided anything except that teachers need to be intelligent, knowledgeable, interested in their work, and sympathetic towards children? The question is justified.

V. A new approach

A comparatively recent development is the Teacher Characteristics Study sponsored by the American Council on Education. This has as its two main objectives the identification and analysis of patterns of teacher-behaviour, and the development of materials useful for its prediction.

This study has involved extensive observation of teachers with a view to understanding significant patterns of their behaviour and of the responses of the pupils. Reports of especially effective and ineffective behaviour on the part of teachers were collected and analysed and from these a list of critical requirements was developed. A critical requirement was defined as "any observable teacher behaviour which might make the difference between success and failure in teaching". Those aspects of teaching carried on satisfactorily by almost all teachers were omitted, and only those that seemed to differentiate between extremes of efficiency were included. This method is called the Critical Incident Technique.

The critical requirements were found to be classifiable under three main headings.

1. *Personal qualities*
   (a) Optimism.
   (b) Fairness.
   (c) Self-control.
2. *Professional qualities*
   (a) Knowledge of subject matter and techniques of teaching.
   (b) Ability to get student response.
   (c) Business-like approach.
3. *Social qualities*
   (a) Sympathetic, understanding.
   (b) Democratic.
   (c) Friendly, commending.
   (d) Ability to judge reactions of others.

As a result of this analysis an Observation Blank was constructed and used for making objective records in other studies. These were concerned with the differences between the behaviour of elementary and secondary school teachers, the attitudes and educational viewpoints of teachers, and the relationship between professional and personal data and the judged effectiveness of the teacher.

This method of studying teachers at work rather than attempting to assess their efficiency seems to have useful possibilities. It should lead to a better understanding of what a teacher actually does and of its effects on the children, and so to a clearer realisation of the qualities needed in teachers and of the tasks to be undertaken in training them. It also suggests a method of judging the value of a teacher's activities and of predicting their probable outcomes.

VI. Conclusions

As the reader will have realised, there is a great deal of information available about teachers and teaching efficiency. Some of it is contradictory and some of it is inconclusive. Differences between results obtained appear to be more marked than similarities, and correlations between teaching efficiency and other qualities of the teacher are usually small.

The most obvious deduction from the available evidence is that there is no one pattern of successful teacher. Teaching is a complex process calling for many different abilities, no one of which is by itself sufficient to ensure success. At the same time, each of these abilities makes a small but necessary contribution to success. Limited success results where some but not all of these abilities are exercised, failure is more than probable where too many of them are lacking. An adequate level of health, intelligence and knowledge, the desire to teach, a liking for children and the ability and will to use and apply knowledge in their education are major factors contributing to success in teaching. If to these can be added a pleasing personality and a lively interest in the world around, the result is likely to be a very good teacher indeed.

It is noticeable that there are considerable differences between the students selected by different training colleges, although the basic qualifications for entry are the same for all. Selection is made in accordance with the assumptions of the selectors about what is desirable teaching. Some emphasise one aspect of the work, some emphasise others, and they choose their students accordingly. This difference of emphasis is sometimes made because of differences in the age groups or of the special subjects to be taught. It is arguable that since all teachers are not going to be employed on exactly the same work, then they should not all be the same kind of person. Training is based on the same kind of assumption as governs selection, and the results are equally good for different colleges.

The principle here is that the colleges select students who, they expect, will benefit from the training they can provide. The principle is a sound one, and it points to the direction which future research should take. Much general research on the concomitants of teaching efficiency has been carried out. In the light of the knowledge gained, the most useful contribution which the colleges can now make would be in the form of studies of the success of their own students in relation to the factors taken into account in selecting them and the methods used in training them. In this way individual colleges could do much to raise their own already high standards, and add to our store of knowledge of factors contributing to success in teaching. Some work of this kind is being done.

A second line of study might be concerned with factors which contribute to successful teaching of particular age groups and subjects, and in particular schools. Detailed knowledge of this kind, rather than general and vague information, would be invaluable in fitting potential teachers into the kind of work where they could make the best use of their talents. There is a great deal of justification for thinking that success in teaching is often a matter of finding the right niche.

This is not to say that there are no people who would not always and anywhere be failures as teachers. Such people do exist, but many of them have no wish to become teachers. Some do, and, in their own interests as well as those of the children, they should not be admitted to the profession. Increased knowledge of what makes a person a good teacher, and of the possibilities of success for different types of people, would prevent much unhappiness for unsuitable applicants for training. There is no profit or happiness for either teacher or children where the former is really unsuited to the work. Human happiness in the classroom may well be one of the best indications of the presence of a successful teacher.

Bibliography

Allard, L. E. (1939). *A Study of the Leisure Activities of Certain Elementary School Teachers of Long Island.* New York: Teachers' College, Columbia University.

Arvidson, G. L. (1956). *Some Factors Influencing the Achievement of First Year Secondary Modern School Children.* Unpublished Ph.D. Thesis, University of London Library.

Austin, F. M. (1931). "An analysis of the motives of adolescents for the choice of the teaching profession." *Brit. J. Educ. Psychol.*, 1, 87-103.

Barr, A. S. (1948). "The measurement and prediction of teaching efficiency: A summary of investigations." *J. Exp. Educ.*, 16, 203-208.

Barr, A. S., *et al.* (1952). Report of the Committee on the Criteria of Teacher Effectiveness. *Rev. Educ. Res.*, 22, 238-263.

Bent, R. K. (1937). "Relationships between qualifying examinations, various other factors, and student teaching performance at the University of Minnesota." *J. Exp. Educ.*, 5, 251-255.

Birkinshaw, M. (1935). *The Successful Teacher.* London: Hogarth Press.

Bliss, W. B. (1922). "How much mental ability does a teacher need?" *J. Ed. Res.*, 6, 33-41.

Blum, M. L. (1936). "An investigation of the relation existing between students' grades and their ratings of the instructor's ability to teach." *J. Educ. Psychol.*, 27, 217-221.

Boardman, C. W. (1928). *Professional Tests as Measures of Teaching Efficiency in High School.* New York: Teachers' College, Columbia University.

Boyce, A. C. (1912). "Qualities of merit in secondary school teachers." *J. Educ. Psychol.*, 3, 144-157.

Boyce, R. B., and Bryan, R. C. (1944). "To what extent do pupils' opinions of teachers change in later years?" *J. Educ. Res.*, 37, 698-705.

Brookover, W. B. (1945). "The relation of social factors to teaching ability." *J. Exp. Educ.*, 13, 191-205.

Brown, E. E. (1938). *The Selection and Education of Oklahoma High School Teachers.* Oklahoma: Harlow Publishing Co.

Bryan, R. C. (1937). *Pupil Rating of Secondary School Teachers.* New York: Teachers' College, Columbia University.

Buckingham, B. R. (1920). "A proposed index of efficiency in teaching United States History." *J. Educ. Res.*, 1, 161-171.

Buckingham, B. R. (1923). "An unrecognised ability in teaching." *J. Educ. Res.*, 7, 59-62.

Cattell, R. B. (1931). "The assessment of teaching ability." *Brit. J. Educ. Psychol.*, 1, 48-72.

Charters, W. W., and Waples, D. (1929). *The Commonwealth Teacher-Training Study.* Chicago: Univ. Chicago Press.

Connor, W. L. (1920). "A new method of rating teachers." *J. Ed. Res.*, 1, 338-358.

Cook, W. W. (1942). "Predicting success of graduate students in a College of Education." *School and Society*, 56, 192-195.

Cook, W. W., and Leeds, C. H. (1947). "Measuring the teaching personality." *Educ. and Psychol. Meas.*, 7, 399-410.

Cook, W. W., Leeds, C. H., and Callis, R. (1951). *Minnesota Teacher Attitude Inventory.* New York: Psychological Corporation.

Crabbs, L. M. (1925). *Measuring Efficiency in Supervision and Teaching.* New York: Teachers' College, Columbia University.

Dolch, E. W. (1920). "Pupils' judgments of their teachers." *Ped. Sem.*, 27, 195-199.

Domas, S. J. (1950). *Report of an Exploratory Study of Teacher Competence.* Cambridge, Mass.: New England School Development Council.

Domas, S. J., and Tiedeman, D. V. (1950). "Teacher competence: An annotated bibliography." *J. Exp. Educ.*, 19, 101-218.

Drucker, A. J., and Remmers, H. H. (1951). "Do alumni and students differ in their attitudes towards instructors?" *J. Educ. Psychol.*, 42, 129-143.

Eliassen, R. H., and Martin, R. L. (1945). "Pretraining selection of teachers during 1940-43." *J. Educ. Res.*, 38, 666-677.

Evans, K. M. (1946). *A Study of Attitude towards Teaching as a Career.* Unpublished M.A. Thesis, University of London Library.

Evans, K. M. (1951). "A critical survey of methods of assessing teaching ability." *Brit. J. Educ. Psychol.*, 21, 89-95.

Evans, K. M. (1952a). *A Study of Teaching Ability at the Training College Stage in Relation to the Personality and Attitudes of the Student.* Unpublished Ph.D. Thesis, University of London Library.

Evans, K. M. (1952b). "A study of attitude towards teaching as a career." *Brit. J. Educ. Psychol.*, 22, 63-70.

Evans, K. M. (1953). "A further study of attitude towards teaching as a career." *Brit. J. Educ. Psychol.*, 23, 58-63.

Evans, K. M. (1957). "Is the concept of 'interest' of significance to success in a teacher training course?" *Educ. Rev.*, 9, 205-211.

Evans, K. M. (1958). "An examination of the Minnesota Teacher Attitude Inventory." *Brit. J. Educ. Psychol.*, 28, 253-257.

Freyd, M. (1923). "A graphic rating scale for teachers." *J. Educ. Res.*, 8, 433-439.

Geyer, D. L. (1946). "Qualities desired in college instructors." *School and Society*, 63, 270-271.

Gotham, R. E. (1945). "Personality and teaching efficiency." *J. Exp. Educ.*, 14, 157-165.

Grieder, C., and Newburn, H. K. (1942). "Temperament in prospective teachers." *J. Educ. Res.*, 35, 683-693.

Hampton, N. D. (1951). "An analysis of supervisory ratings of elementary teachers graduated from Iowa State Teachers' College." *J. Exp. Educ.*, 20, 179-216.

Heilman, J. D., and Armentrout, W. D. (1936). "The rating of college teachers on ten traits by their students." *J. Educ. Psychol.*, 27, 197-216.

Hellfritzch, A. G. (1945). "A factor analysis of teacher abilities." *J. Exp. Educ.*, 14, 166-199.

Henrikson, E. H. (1943). "Comparisons of voice and teaching ability." *J. Educ. Psychol.*, 34, 121-123.

Herda, F. J. (1935). "Some aspects of the relative instructional efficiency of men and women teachers." *J. Educ. Res.*, 29, 196-203.

Hollis, A. W. (1935). *The Personal Relationship in Teaching.* Unpublished M.A. Thesis, University of Birmingham Library.

Hult, E. (1945). "Study of achievement in educational psychology." *J. Exp. Educ.*, 13, 174-190.

Jayne, C. D. (1945). "A study of the relationship between teaching procedures and educational outcomes." *J. Exp. Educ.*, 14, 101-134.

Jensen, A. C. (1951). "Determining critical requirements for teachers." *J. Exp. Educ.*, 20, 79-85.

Jersild, A. T. (1940). "Characteristics of teachers who are 'liked best' and 'disliked most'." *J. Exp. Educ.*, 9, 139-151.

Johnson, E. C., and Morris, E. H. (1937). "Considerations concerning the selection of prospective teachers." *School and Society*, 46, 222-224.

Jones, R. D. (1946). "The prediction of teaching efficiency from objective measures." *J. Exp. Educ.*, 85-99.

Kent, R. A. (1920). "What should teacher-rating scales seek to measure?" *J. Educ. Res.*, 2, 802-807.

Knight, F. B. (1922a). *Qualities Related to Success in Teaching.* New York: Teachers' College, Columbia University.

Knight, F. B. (1922b). "Qualities related to success in elementary school teaching." *J. Educ. Res.*, 5, 207-216.

Knight, F. B. (1923). "The effect of the 'acquaintance factor' upon personal judgments." *J. Educ. Psychol.* 14, 129-142.

Kratz, H. E. (1896). "Characteristics of the best teacher as recognized by children." *Ped. Sem.*, 3, 413-418.

LaDuke, C. V. (1945). "The measurement of teaching ability." *J. Exp. Educ.*, 14, 75-100.

Lancelot, W. H. (1935). "A study of teaching efficiency as indicated by certain permanent outcomes," in Walker, H. M. (Ed.), *The Measurement of Teaching Efficiency.* New York: Macmillan Co.

Lawton, J. (1939). "A study of factors useful in choosing candidates for the teaching profession." *Brit. J. Educ. Psychol.*, 9, 131-144.

Leeds, C. H. (1950). "A scale for measuring teacher-pupil attitudes and teacher-pupil rapport." *Psychological Monographs: General and Applied*, No. 312.

Lins, L. J. (1946). "The prediction of teaching efficiency." *J. Exp. Educ.*, 15, 2-60.

Lovell, K. (1951). *An Investigation into Factors Underlying Teaching Ability in Primary and Secondary Modern Schools with a view to Improving the Methods of Selecting Potential Teachers.* Unpublished M.A. Thesis, University of London Library.

MacLear, M. (1913). "The teacher's need of a community life." *Ped. Sem.*, 20, 539-541.

Martindale, F. E. (1951). "Situational factors in teacher placement and success." *J. Exp. Educ.*, 20, 121-178.

Mason, F. V. (1930). "A study of seven hundred maladjusted school teachers." *Mental Hygiene*, 15, 576-599 (Cited by Allard, *loc. cit.*).

Mathews, L. H. (1940). "An item analysis of measures of teaching ability." *J. Educ. Res.*, 33, 576-580.

Mead, A. R. (1929). "Qualities of merit in good and poor teachers." *J. Educ. Res.*, 20, 239-259.

Mead, A. R., and Holley, C. E. (1916). "Forecasting success in practice teaching." *J. Educ. Psychol.*, 7, 495-497.

Morris, E. H. (1929). *Personal Traits and Success in Teaching.* New York: Teachers' College, Columbia University.

Nandi, T. K. (1940). *An investigation into the assessment of teaching ability.* Unpublished B.Ed. Thesis, University of Edinburgh Library. Summary in *Brit. J. Educ. Psychol.*, 11, 62-63.

Odenweller, A. L. (1936). *Predicting the Quality of Teaching.* New York: Teachers' College, Columbia University.

Pace, C. R. (1950). " Opinion and action: A study in the validity of attitude measurement." *Educ. and Psychol. Meas.*, 411-419.

Panton, J. H. (1934). *The Assessment of Teaching Ability with Special Reference to Men Students in Training.* Unpublished M.A. Thesis, University of London Library.

Payne, E. G. (1918). " Scholarship and success in teaching." *J. Educ. Psychol.*, 9, 217-219.

Phillips, W. S., and Greene, J. E. (1939). " A preliminary study of the relationship of age, hobbies and civil status to neuroticism among women teachers." *J. Educ. Res.*, 30, 440-444.

Pinsent, A. (1933). " Pre-college teaching experience and other factors in the teaching success of university students." *Brit. J. Educ. Psychol.*, 3, 109-126, and 201-221.

Porter, W. A. (1942). " Pupil evaluation of practice teaching." *J. Educ. Res.*, 35, 700-704.

Retan, G. A. (1943). " Emotional instability and teaching success." *J. Educ. Res.*, 37, 135-141.

Reymert, M. L. (1917). " The psychology of the teacher: An introductory study." *Ped. Sem.*, 24, 521-558.

Rinsland, M. A. O. (1938). " A test for measuring teachers' knowledge of the conduct and personality of children from six to eight years of age." *J. Exp. Educ.*, 6, 307-317.

Rolfe, J. F. (1945). " The measurement of teaching ability." *J. Exp. Educ.*, 14, 52-74.

Rostker, L. E. (1945). " The measurement of teaching ability." *J. Exp. Educ.*, 14, 6-51.

Ruediger, W. C., and Strayer, G. D. (1910). " The qualities of merit in teachers." *J. Educ. Psychol.*, 1, 272-278.

Ryans, D. G. (1951). " A study of the extent of association of certain professional and personal data with judged effectiveness of teacher behaviour." *J. Exp. Educ.*, 20, 67-77.

Ryans, D. G. (1952). " A study of criterion data (A factor analysis of teacher behaviors in the elementary school)." *Educ. and Psychol. Meas.*, 12, 333-344.

Ryans, D. G., and Wandt, E. (1952a). " Investigations of personal and social characteristics of teachers." *J. Teacher Education*, Sept.

Ryans, D. G., and Wandt, E. (1952b). " A factor analysis of observed teacher behaviors in the secondary school." *Educ. and Psychol. Meas.*, 12, 574-586.

Sandiford, P., *et al.* (1937). *Forecasting Teaching Ability.* Toronto: Dept. of Educational Research, Univ. of Toronto.

Sanford, C. W., and Trump, J. E. (1950). " Teacher Education—IV. Preservice Selection," in Monroe, W. S. (Ed.), *Encyclopaedia of Educational Research*, 1390-1396. New York: Macmillan Co.

Schwartz, A. N. (1950). " A study of the discriminating efficiency of certain tests of the primary source traits of teachers." *J. Exp. Educ.*, 19, 63-94.

Seagoe, M. V. (1945a). " Permanence of interest in teaching." *J. Educ. Res.*, 38, 678-684.

Seagoe, M. V. (1945b). " Prognostic tests and teaching success." *J. Educ. Res.*, 38, 685-690.

Seagoe, M. V. (1946). " Prediction of in-service success in teaching." *J. Educ. Res.*, 39, 658-663.

Shannon, J. R. (1936). " A comparison of three means for measuring efficiency in teaching." *J. Educ. Res.*, 29, 501-508.

Skinner, W. A. (1949). *An Investigation into Assessment of Teaching Ability in Teachers of Technical Subjects.* Unpublished M.A. Thesis, University of London Library.

Smalzried, N. T., and Remmers, H. H. (1943). " A factor analysis of the Purdue Rating Scale for Instructors." *J. Educ. Psychol.*, 34, 363-367.

Sprague, H. A. (1917). " Score-card for rating student-teachers in training and practice." *Ped. Sem.*, 24, 72-80.

Stott, M. B. (1950). " What is occupational success? " *Occupational Psychology*, 24, 105-112.

Stuit, D. B. (1937). " Scholarship as a factor in teaching success." *School and Society*, 46, 382-384.

Symonds, P. M. (1950). " Reflections on observations of teachers." *J. Educ. Res.*, 43, 688-696.

Thomson, G. H. (1921). " A rating scale for teaching ability in students." *J. Exp. Ped.*, 6.

Torgerson, T. L. (1937). " The measurement and prediction of teaching ability." *Rev. Educ. Res.*, 7, 242-246.

Tudhope, W. B. (1942 and 1943). " A study of the training college final teaching mark as a criterion of future success in the teaching profession." *Brit. J. Educ. Psychol.*, 12, 167-171, and 13, 16-23.

Tudhope, W. B. (1944). " Motives for the choice of the teaching profession by training college students." *Brit. J. Educ. Psychol.*, 14, 129-141.

Upshall, C. C. (1942). " The validity of composite faculty judgment as a method of identifying undesirable prospective elementary school teachers." *J. Educ. Res.*, 35, 694-699.

Valentine, C. W. (1934). " An enquiry as to the choice of the teaching profession by university students." *Brit. J. Educ. Psychol.*, 4, 237-259.

Vernon, P. E. (1939). " Educational abilities of training college students." *Brit. J. Educ. Psychol.*, 9, 233-250.

Vernon, P. E. (1953). " The Psychological Traits of Teachers," in Hall, R. K., Hans, N., and Lauwerys, J. A. (Ed.), *The Year Book of Education.* London: Evans Bros.

Von Haden, H. I. (1946). " An evaluation of certain types of personal data employed in the prediction of teaching efficiency." *J. Exp. Educ.*, 15, 61-84.

Wall, W. D. (1955). *Education and Mental Health.* London: UNESCO/Harrap.

Ward, L. B., and Kirk, S. A. (1942). " Studies in the selection of students for a teacher's college." *J. Educ. Res.*, 35, 665-672.

Whitney, F. L. (1923). " The analysis of teaching functions." *J. Educ. Res.*, 297-308.

Witham, E. C. (1914). " School and teacher measurement." *J. Educ. Psychol.*, 5, 267-278.

Yeager, T. C. (1935). *An Analysis of Certain Traits of High School Seniors Interested in Teaching.* New York: Teachers' College, Columbia University.

Educational Guidance—I

GUIDANCE SERVICES IN THE UNITED STATES

by Alfred Yates

THE editors of the Year Book of Education for 1955, which was devoted to a study of the processes involved in guidance and counselling, succeeded, after a few false starts, in framing a definition of guidance which was accepted by the contributors to that volume. They defined it as the ' process of helping individuals through their own efforts to discover and develop their potentialities both for personal happiness and social usefulness '. Even when qualified by the epithet ' educational ' and restricted to the activities of those responsible for the welfare of pupils in schools and other scholastic institutions, this definition still suggests a function far broader in scope than the one that has been traditionally accorded to teachers. The function may be regarded as that of an honest broker, mediating between, on the one hand, the characteristics, potentialities and aspirations of the pupils and, on the other, the opportunities that the schools can afford.

The purpose of this series of articles is to report the evidence that has been collected concerning the effectiveness of the various processes involved in educational guidance, as defined above. Such a review, it is hoped, will enable those responsible for them to decide to what extent the present arrangements made in our schools for helping individuals to discover and develop their potentialities are adequate.

This introductory article will be concerned, in the main, with the kinds of guidance and counselling services that are provided in many of the school systems in the United States. There have been a number of empirical studies designed to evaluate these services. What follows is a summary of the results of these studies and of the views of those who have taken part in guidance services and who have submitted their experience to critical analysis. Before furnishing this summary, however, it is necessary to give a brief outline of the kind of guidance service that is to be found in a typical American School District.

Guidance and Counselling Services in the United States

Full-time professionally-trained counsellors are employed in many American school systems. They are people who have undertaken courses in educational psychology and mental measurement and who are responsible for directing and co-ordinating the educational guidance of all the pupils in the schools with which they are concerned. It is not uncommon to find that one counsellor is responsible for a group of elementary schools in a district, and is assisted by ' teacher-counsellors ' working within the schools themselves. The ' teacher-counsellers ' are relieved of some of their teaching duties in order to serve in this capacity. The larger elementary schools, however, and most of the high schools have one or more full-time counsellors on the staff as well as the assistance of a number of teacher-counsellors. The following account provided by Traxler of part of the guidance programme in a high school shows clearly why additional man-power is needed.

Every pupil who enters the Plainfield High School in New Jersey, U.S.A., is required to answer a questionnaire consisting of more than 150 items. Each response is entered on the pupil's record sheet which is 24½ inches long, 11 inches wide, and affords space for over a thousand entries. These record sheets have been adapted to a visible filing system to facilitate their rapid and accurate use by administrative officers and teachers and are housed in the office of the school's Director of Tests and Records.

The school has a test scoring machine which makes it possible for the results of tests to be made rapidly available—usually on the day they are administered. During the course of one academic year 11,080 tests were administered to the 1,820 pupils as part of the annual testing programme. These tests were additional to those constructed and administered by the teachers themselves, which can also be scored on the machine. These are produced at the rate of about thirty a year and are added to the school's permanent library of tests. Teachers may draw on the stocks of this library, which contains over 18,000 tests, at any time during the year when they deem it necessary to evaluate the progress made by their pupils or when they require tests for diagnostic purposes.

These arrangements, it is claimed, make it possible for teachers and counsellors to chart in detail the progress of each pupil. The wealth of information derived from the annual testing programme and from the teachers' tests and examinations demonstrates the emergent pattern of each child's abilities and aptitudes. This information can be related to the facts about the child's background, interests, personal qualities and out-of-school experiences obtained from his questionnaire responses and by means of subsequent observations and assessments.

The evaluation of programmes of educational guidance

Programmes of educational guidance of the kind described above, professionally directed, based on objective measurements of abilities and aptitudes, on cumulative records, and on regular counselling interviews in which teachers and parents, as well as the pupils concerned often take part, have been evaluated in a number of follow-up studies.

Some of these studies are listed in the bibliography that follows this article. Although they exhibit some differences in experimental design and with regard to the criteria employed in evaluating the effectiveness of the guidance programmes under review, the majority of them involve a comparison between the progress made during school and afterwards by a group of guided pupils and that made by a matched group of pupils who were not afforded guidance. A brief account of the study reported by Rothney and Roens may, therefore, be regarded as fairly typical of the follow-up studies that have been carried out.

This inquiry involved three junior high schools and one senior high school in Arlington, near Boston. The pupils in the three junior high schools followed a common curriculum up to the seventh grade (age 13); during their eighth grade they were free to add certain elective subjects—Latin, business practice, and extra science, for example; and in the ninth grade they joined one of a number of courses, very roughly corresponding to our own grammar, modern, technical and commercial courses. These courses were continued in the senior high school, and, indeed, the work in the ninth grade was regarded as a means of effecting a smooth transition from one school to the next in that the pupils chose the course they wished to follow only after visiting the senior high school, becoming acquainted with its programme of studies and, in some instances, after availing themselves of the senior high school's counselling services.

For the purposes of this follow-up study, two groups were formed; each containing 130 pupils. Those in the experimental group were nominated by their schools. This group contained children who were in need of guidance because they had encountered difficulties of some kind and did not appear to be progressing as satisfactorily as they might be. It also contained children selected because they were unusually talented and required special attention to ensure that their superior potentialities were fully realised.

The children in the control group were selected by the experimenters who endeavoured to match each child in the experimental group with one who was similar with respect to the following: sex; chronological age; mental age; grade; level of attainments; religion; father's occupation; socio-economic status.

The children in the control group were afforded only the rudimentary form of guidance that was provided in the schools at the time of the experiment. Those in the experimental group were given systematic and continuous guidance, supervised by trained counsellors and involving regular discussions in which the pupil's teachers and parents also took part. A testing programme was devised for them comprising regular assessments of general, mechanical, spatial and clerical abilities and attainments in the basic subjects—reading, vocabulary, mathematics, etc. Their school examination marks were also recorded, along with the test scores, on cumulative record cards. The counsellor's interpretation of the evidence collected in this way about each pupil's progress was discussed with the teachers concerned and suggestions were made for remedial education, if this was found to be necessary, and for the adaptation of the curriculum and methods of teaching to suit each child's particular needs. The counselling services provided also had a therapeutic element in that the counsellors were able to discuss a child's personal problems with the child himself or with his teachers and parents, and to suggest ways in which his difficulties might be overcome. It was part of the policy of the counsellors who took part in this inquiry to leave pupils and their parents free to select such courses as they deemed suitable, the counsellors acting merely as advisers who brought forward what they considered to be relevant evidence which would enable the people concerned to make a wise choice. A child was advised against embarking on a particular course or subject only if the odds seemed to be overwhelmingly against his being able to succeed in it. Before this form of guidance was introduced the distinctions between the courses provided in the schools had been fairly rigidly maintained. The counsellors were able, however, to persuade the schools to provide ' cross-curricular ' courses for about a fifth of the experimental group.

A composite criterion was used to evaluate the effectiveness of the guidance provided for the experimental group in this inquiry. The records of each child's progress during the course of the experiment provided a basis of comparison between the experimental and control group, particularly in terms of their academic successes. Two sets of reports were also obtained from the pupils themselves. The first of these was obtained shortly before they graduated from the high school, the second eighteen months afterwards. These reports were in the form of answers to questionnaires designed to assess the extent to which the pupils were able to make satisfactory adjustments to life in college or in whatever job they took up after leaving school. The responses were obtained individually and supplemented by an interview. The scoring of the responses and the judgments concerning whether or not a particular response indicated superior adjustment were made the responsibility of independent assessors who had not taken part in the inquiry, and who did not know to which group—experimental or control—any individual had belonged.

This inquiry has been described in detail because it is fairly typical of a number of such studies. Some of the other inquiries listed in the bibliography also included as part of the criterion for assessing the value of guidance services the opinions expressed about these services by those pupils who had made use of them. In those circumstances in which the use of the guidance and counselling services is largely voluntary—as it is for the students in many colleges and universities, for example—this criterion has a limited value in indicating the kinds of students who avail themselves of the opportunities provided. The fact, however, that a sizeable percentage of those who do so, tend to testify to the benefits they receive from it is not very impressive evidence. As one commentator has observed: " The trouble is that the same sort of evidence can be obtained for the success of fortune-tellers, phrenologists, and faith-healers " (Tyler). The following summary of results, therefore, is confined to those studies in which the criteria employed consisted either of objective measures of the subsequent progress and achievements of the pupils concerned or of controlled observations and assessments made by independent judges, as described above.

The results of follow-up studies of guidance programmes

The following results have been reported in one or more of the studies listed in the bibliography at the end of this article. Only those results have been quoted which have either been obtained independently by several researchers or, if derived from a single inquiry, are based on substantial evidence.

Pupils who enjoyed the benefits of guidance and counselling services have been shown to be superior to control groups of pupils who were not afforded these benefits in the following respects:

(*a*) their average level of attainments was significantly higher;

(*b*) they made fewer changes of curriculum, indicating presumably that their choices of subjects and courses were more suitable;

(*c*) a significantly higher proportion of them proceeded to colleges, universities or to other forms of higher education;

(*d*) the average I.Q. of those who proceeded to colleges and universities was more in line with academic and professional requirements than that of the pupils from the control groups—in other words, their choices in this respect were more realistic;

(*e*) those who left school to go into employment were judged by independent assessors to have adopted more realistic attitudes towards their future vocations, to be better adjusted to their environment and a significantly greater proportion of them had undertaken part-time courses designed to further their vocational plans and were pursuing these courses successfully;

(*f*) there were fewer delinquents among the guided pupils, more of them belonged to clubs and other social organisations and they were generally rated as being emotionally more stable and mature than the members of the control groups.

Current views on educational guidance in the United States

The majority of American educationists appear to be satisfied that a case has been made out for the effectiveness of the kinds of guidance and counselling services that have been operating for many years in various parts of the United States. Their belief in the value of these services does not rest solely on the evidence of the empirical studies summarised above, but on the accumulated experience of those who have participated in guidance and counselling services. There is now a considerable measure of agreement not only that educational guidance should be afforded to all pupils, but also about how it can best be organised.

The role of the teacher

Some years ago the trend appeared to be towards increasing the number of full-time counsellors and ultimately towards placing the responsibility for educational guidance entirely in their hands. The arguments put forward by those who pressed for the ' complete professionalisation ' of the guidance service were based largely on the inability of teachers to perform the function for themselves. Educational guidance, it was argued, involving as it does the assessment of children's abilities, attainments, aptitudes, attitudes and interests by means of standardised tests and other instruments which can only be effectively employed by those who have undergone special training for the purpose, lies outside the competence of the average teacher. Furthermore, it was claimed that the traditional conception of the relationship between teachers and pupils—the view that teachers are in daily contact with their pupils and therefore have opportunities to observe their behaviour and to form judgments about their educational needs that might well be as valid or more so than those based on the children's responses to test or questionnaire items—no longer applied in the majority of secondary schools. The tendency towards departmentalisation and specialisation resulted, it was claimed, in teachers and children coming together for only isolated periods. In these circumstances no teacher is able to become intimately acquainted with a particular group of children. The professional counsellor, therefore, is not at any disadvantage, compared with the teachers, as far as opportunities for intimate knowledge of the children are concerned, and, indeed, being relieved of normal teaching duties, having access to the collated information about each child's test scores and personal data, and being free to interview any particular child when the need arises, he is, on the whole, in a better position than any individual teacher to assess a child's needs and potentialities and to advise him about his educational and vocational plans.

There are now signs that opinion, amongst those who are concerned with the organisation of educational guidance in the United States, is moving towards placing greater emphasis on the role of the teacher. This does not mean that it is proposed to dispense with the services of the full-time professional counsellors. There would appear to be fairly general agreement that their services are invaluable. What has come to be recognised, however, is that a system of educational guidance based entirely on the activities of professional counsellors is impracticable. Experience has shown that even a group of counsellors cannot adequately cater for the needs of every child in a school. It is also clear that unless the teachers in a school are co-operating wholeheartedly with the counsellors, not merely to the extent of supplying the information required for the children's cumulative records, but by actively participating in the task of trying to discover what each child's educational needs are, the resultant services are less than satisfactory.

The objection that teachers in secondary schools, because of the tendency of each member of the staff to be a specialist and therefore to have little opportunity for making regular contacts with any specific group of children, has been overcome in some schools by the use of some members of the staff as general practitioners. These teachers, responsible for what is called in some schools the ' basic course programme ' teach a number of subjects to children of a particular grade and are given special responsibility, in co-operation with the counsellor for the guidance of these children.

The present consensus of opinion in the United States would seem to be that educational guidance must involve the teachers but that it is more effectively carried out as a team effort, the participants being the teachers, the teacher-counsellors (the teachers responsible for the ' basic course programme ' who are freed from teaching duties for a part of each day) and the professional counsellor who serves to co-ordinate the efforts of the team and is responsible for maintaining the necessary records about each child.

Continuous guidance for every pupil

There is general agreement, too, that the favoured treatment usually accorded to pupils suffering from some disability or abnormality should, as far as possible, become the prerogative of every child. In this country child guidance has been conceived as a service given to children who are maladjusted or in need of remedial educational treatment. It is now recognised that in order to be able to make the best possible use of his talents every child requires from time to time help in overcoming the obstacles that impede his educational progress. Such help can only be timely and effective if the child is supervised by teachers who are able to give close and regular attention to his progress, who are able to select from a wide variety of texts and teaching materials those which are suited to his level of attainment and rate of progress, and who have available to them sufficient information about the child's characteristics to enable them to diagnose the cause of difficulties that he encounters or, better, to anticipate them.

This ideal programme suggests that to be fully effective educational guidance must not only be available to every child but must be a continuous process. Episodic acts of guidance offered at crucial points in a child's career are clearly of limited value. Indeed, such guidance may be nugatory as when, for example, it is based on the result of an examination which may be wholly unrepresentative of the child's normal level of performance.

Child-centred guidance

The need is also emphasised for what is sometimes described as 'child-centred' guidance or counselling. There is a recognised risk of guidance becoming something akin to direction. For example, if a pupil has to be advised concerning which of a number of possible courses he should attempt in a secondary school, the teacher who has taken the trouble systematically to analyse the child's abilities and aptitudes and has reached the conclusion that one particular course is ideally suited to his capacities may be tempted to press his case too strongly. Admittedly, it would be irresponsible of a teacher to allow a child to make an unrealistic choice, involving the strong likelihood of subsequent failure and frustration. On the other hand, a child who is denied a preferred course may, for that reason, be inadequately motivated to succeed in any alternative that he has been persuaded to accept.

Clearly, educational guidance calls for tact and nice judgment as well as a skilful assessment of a child's needs and capabilities. Its ultimate aim, moreover, is recognised as being to develop in each child the necessary insights that will enable him eventually to undertake the sensible direction of his own affairs.

Educational guidance in this country

The foregoing outline of the guidance and counselling services in the United States and of the conclusions of those who have reviewed their operation is intended to provide a background against which our own somewhat different forms of educational guidance can be examined. The arrangements made for the guidance of pupils in the schools in this country certainly seem to be far less elaborate and comprehensive than those in American schools. Such a comparison, however, is not altogether valid. In this country a major act of educational guidance takes place outside the schools, so to speak. The procedures for allocating primary school leavers to different secondary schools or courses are designed for some of the same purposes as the guidance services in an American high school. In many of the latter, one result of the counsellor's assessments is that a pupil is advised to enter a college preparatory course, a technical course or a general course. If the term 'grammar' is equated with the first of these and that of 'modern' can be equated with the last, it is perhaps arguable that the differences are to be found, at least in part, in the techniques employed rather than in the ends served by the two forms of guidance. Before and after the eleven-plus examination, children in many of our schools are also 'streamed'. Further guidance is afforded in the secondary schools insofar as provision is made for the children to follow different curricula. In grammar schools, for example, some children may be allowed to tackle a second foreign language and some choice is afforded of subjects to be taken up to G.C.E. levels. In many other types of secondary school, alternative courses are nowadays offered, particularly for the senior pupils, and the allocation of pupils to these courses is, for the most part, based on assessments of the children's progress, levels of attainment, interests and vocational ambitions.

Nevertheless, although the contrast between the two systems may not be quite so pointed as it first appears, there is little doubt that the transatlantic concept of educational guidance has not yet been fully accepted in this country. This concept, as was shown earlier, involves a continuous process of guidance, supervised if not administered by professionally-trained counsellors, based on a wealth of objective test data, on cumulative records, and on the close observation of each child's personal characteristics. It involves, too, the readiness to adjust curricula and teaching methods to suit as far as is possible the ascertained needs of each individual pupil. Even if the findings which were quoted earlier are accepted as valid evidence of the effectiveness of programmes of educational guidance developed along these lines they could only be accorded the status of hypotheses as far as our own educational system is concerned. There are so many differences in aim and organisation between the schools in the two countries that practices which are demonstrably appropriate in one are not necessarily suitable for the other.

The extent, therefore, to which our present forms of educational guidance are effective as a means of enabling the pupils in our schools to ' discover and develop their potentialities for personal happiness and social usefulness ' cannot be determined by reference to the evidence of research and inquiry into the operation of guidance services in the United States, although this evidence is useful in suggesting the kinds of experiment that might be usefully tried out in our schools.

The need for improved methods of educational guidance

The evidence, cited in this article, in support of the claim that systematic guidance has assisted the educational progress of pupils in American schools provides perhaps a sufficient justification for exploring the possibilities of securing improvements along comparable lines in our own schools. There are also additional reasons for reviewing our procedures. However satisfactory these may have been in the past, there have clearly been important changes in recent years which have affected both sides of the equation with which educational guidance is concerned: our knowledge of the characteristics and educational needs of schoolchildren; and the demands of and opportunities afforded by the adult society for which they are being prepared.

Studies of children's development

The systematic studies of children's intellectual and social development, which belong largely to this century, have yielded information which requires to be interpreted in relation to the kind of guidance that should be offered to the children in our schools. The inquiries conducted by Piaget and his associates into the development of children's thinking and, perhaps more particularly, the investigations that have been prompted in this and other countries by his pioneering efforts are producing results that may well call for a revision of our ideas concerning the appropriate content and sequence of topics in the curricula—particularly those in science and mathematics—in both primary and secondary schools.

During the last thirty years considerable ingenuity has been employed in the construction of tests of various kinds and many careful analyses have been made of the results obtained by administering batteries of these tests to children of different ages.

The task of interpreting the evidence thus obtained concerning the nature and growth of children's abilities is by no means complete. In this connexion it is relevant also to quote the work of Olson and his colleagues at the University of Michegan. Their detailed evidence concerning individual differences in rates and pattern of educational progress suggests that the grouping of children in terms of such general measures as mental age or intelligence quotient does not necessarily solve the problem of securing a homogeneous group for teaching purposes. Such, indeed, is the variability in the rhythm of development, not only between one child and another, but during the course of any given child's growth, that they claim that nothing short of individualised instruction—the matching of curriculum and teaching method to the needs of each child—can produce satisfactory results.

These findings call into question the adequacy of class teaching. The probability, even if refined and flexible methods of streaming are used, of assembling in one room forty pupils who are capable of pursuing the same curriculum, of maintaining the same rate of progress and of responding equally to a particular method of instruction would appear to be of a low order.

There has been a growing appreciation of this fact in recent years. In a number of primary schools children are taught in small groups, their assignments of work, the textbooks they use, and the form of instruction they receive being determined by their level of attainment and rate of progress in a particular subject. In some secondary schools, too, class teaching has been abandoned in favour of ' sets ' for such subjects as mathematics and languages in which differences in children's mode and pace of learning are most apparent. It is almost certainly true, moreover, that in many of the schools in which class teaching is still prevalent the teachers cling to this traditional practice not so much because they fail to recognise the need to adapt curricula and methods to the needs of each individual but because the circumstances in which they teach or the facilities with which they are provided compel them to do so. Dividing classes into separate ' sets ' is only possible in a school which enjoys a generous teacher/pupil ratio. Dividing a class into several groups, and giving to each member of such groups the kind of attention he requires, is only feasible if the class is reasonably small and if an adequate supply of suitably graded textbooks and other materials is available.

The demands of contemporary society

The case for reviewing our present arrangement for educational guidance also rests on the new and rapidly changing demands that modern society makes on its members. The Army has often been criticised on the grounds that its training schedules are based on the requirements of the last war but one. Similar criticisms are nowadays levelled against the curricula and methods of the nation's schools, with reference especially to the urgent need not only for a cadre of highly-trained scientists and technologists but for a citizenry favourably disposed to rapid technical change and ready to adapt to its requirements. Educational guidance cannot be exclusively concerned with the assessment of pupils, on the basis of which advice is given concerning which of the available courses of study is likely to be the most beneficial and satisfying for the individual concerned. The possibility must also be envisaged of modifying the kinds of education provided in schools n order best to suit the requirements both of the pupils themselves and of the society of which they form a part. Effective guidance, therefore, may well involve changes in urriculum and adaptations of teaching techniques to ensure that pupils develop those kills that are socially useful. If the social and industrial aims of a society and the echniques it employs for their realisation are undergoing revision, it follows that the preparatory work of the schools should be geared to these changes. This does not imply that the schools should not be concerned with those abiding values on which, even more particularly in times of rapid social and economic change, the stability of society depends. On the other hand, although it is fitting that schools should fulfil their function of transmitting to their pupils the cultural heritage that is their birthright, they are failing in their task if, by maintaining a curriculum that was developed to suit circumstances that no longer obtain, they restrict their pupils' area of choice amongst the wide variety of roles that modern society invites them to fill.

The modifications in curricula and methods of teaching required to meet the needs of a society that is rapidly changing in character, are slowly becoming apparent in secondary schools in various parts of the country. Not only in technical schools as such, but in grammar, modern, and other types of secondary school, courses are now being provided that are very different in content from those that were offered a generation or two ago. Moreover, many pupils now find available to them a choice of courses and there is every indication that, as comprehensive, multilateral and bilateral schools are established or as secondary modern schools develop semi-vocational courses for their more senior pupils, the necessity for such a choice will confront the majority of secondary school pupils at some stage. In these circumstances, there will be a clear need for guidance soundly based on a knowledge of the particular skills and characteristics required for success in the various courses available. Indeed, many teachers in secondary schools, technical colleges and universities are already conscious of this need. Empirical studies are now being made of the problems involved in the selection of students for various university faculties and in the allocation to suitable courses, of students in technical colleges and of pupils in secondary schools.

The increased concern that has recently manifested itself concerning the adequacy of the various procedures that are employed for directing students into suitable courses of secondary and higher education reflects the urgency of modern society's need for adequate supplies of highly-trained individuals. The possibilities of 'wastage' cannot be envisaged with complacency at a time when we are regularly reminded that it is on the developed skill and intellect of our population that our very survival depends. The risk of wastage—in the sense either of directing one person into a course from which he is incapable of deriving benefit or of failing adequately to cultivate the talents of another—must, of course, increase as the number of available courses multiplies, and as their character is modified to meet the needs of a changing society. The risk of wastage has also increased with the change in the social composition of the secondary schools and institutes for further education. The grammar schools, technical colleges and universities now admit in sizeable numbers the children of men and women who in their own day had little or no opportunity to carry their education beyond the elementary stage. These children are largely deprived of those forms of educational guidance which others may expect from their parents. When grammar schools, for example, catered in the main for the sons of those who had themselves attended such schools, and when the range of possible vocations which the boys and their parents would be likely to envisage was limited, the problems of educational and vocational guidance were not difficult to solve. The present situation is vastly different. The task of the grammar schools was formerly the relatively simple one of guiding along well-worn pathways those who had clear ideas about where they were going. Their contemporary problem is how an environment which has undergone rapid and bewildering changes may be interpreted to pupils, many of whom must rely entirely on the staff of the school if they are to choose wisely among the many careers that are now available to them. It is unlikely that this problem will be solved by educational guidance in the narrow sense of measuring abilities and attainments and diverting some pupils into slower moving streams.

Further articles in this series

The evidence and arguments outlined in this introductory article suggest that there is a prima facie case for reviewing our present arrangements for educational guidance. There is a greater variety of courses of secondary and further education available to present-day pupils than was offered a generation ago, and it is almost certain that in the future the problem of mediating between the needs and capacities of children on the one hand and the opportunities afforded by adult society on the other will increase in complexity. It is necessary, therefore, to examine those fields of research which are relevant to this problem. Further articles in this series, therefore, will deal with the following topics: the present state of knowledge concerning the growth and differentiation of children's abilities and aptitudes; the evidence concerning the origin, development and stability of interests; the findings of research into the abilities and characteristics required for success in technical courses; and the extent to which social and emotional factors affect a child's educational progress and the effectiveness in these circumstances of counselling procedures.

References

The following is a list of the books and journal articles to which reference has been made in writing this article.

BLACKWELL, ERNESTINE B., " An Evaluation of the Immediate Effectiveness of the Testing and Guidance Bureau of the University of Texas." *J. Educ. Res.*, 40, 1946.

BRAYFIELD, ARTHUR H. (Ed.), *Readings in Modern Methods of Counseling*. New York: Appleton-Century-Crofts, 1950.

CANTONI, LOUIS J., " Reunion Party: An Unorthodox and Enjoyable Follow-up of the Subjects of a 1939-43 Guidance Experiment." *Clearance House*, 29: April, 1955.

COLE, R. C., " Evaluating a Boys' Club Guidance Program." *Occupations*, 17, 1939.

COLEMAN, WILLIAM, " Some Criteria for Evaluating Elementary School Guidance Services." *Elementary School Journal*, 55: January, 1955.

FROEHLICH, CLIFFORD P., " The Evaluation of Counseling." *Studies in Higher Education, Purdue Univ.*, No. 76, 1951.

FROEHLICH, CLIFFORD P., " Toward More Adequate Criteria of Counseling Effectiveness." *Educ. Psychol. Measmt.*, 9, 1949.

HUTSON, P. W. and WEBSTER, A. D., " An Experiment in the Educational and Vocational Guidance of Tenth-grade Pupils." *Educ. Psychol. Measmt.*, 3, 1943.

JONES, A. J., *Principles of Guidance*. New York: McGraw-Hill, 1945.

KEFAUVER, G. N. and HAND, H. C., " An Appraisal of Guidance." *Occupations*, 12, 1933.

NEWLAND, T. E. and ACKLEY, W. E., " An Experimental Study of the Effect of Educational Guidance on a Selected Group of High School Sophomores." *J. Exp. Ed.*, 5, 1936.

PATERSON, D. G. and CLARK, K. E., " Students' Judgments of Counseling." *J. Higher Educ.*, 14, 1943.

REMMERS, H. H. and WHISLER, L. D., " The Effects of a Guidance Program on Vocational Attitudes." *Studies in Higher Education, Purdue Univ.*, No. 34, 1938.
ROTHNEY, J. W. M. and ROENS, B. A., *Guidance of American Youth*. Cambridge, Mass.: Harvard Univ.
TRAXLER, A. E., *Techniques of Guidance*. New York: Harper, 1945.
TURNEY, AUSTIN H. and MOREHEAD, CHARLES G., " An Experimental Evaluation of a Small High School Counseling Program." *University of Kansas Bulletin of Education*, 8: May, 1954.
TYLER, L. E., *The Work of the Counsellor*. New York: Appleton-Century-Crofts, 1953.
WALL, W. D., " Guidance and Education." *International Review of Education*, II, 3, 1956.
WILLIAMSON, E. G. and BORDIN, E. S., " An Analytical Description of Student Counseling." *Educ. Psychol. Measmt.*, 1, 1941.
WILLIAMSON, E. G. and BORDIN, E. S., " Evaluating Counseling by Means of a Control-Group Experiment." *School and Soc.*, 52, 1940.
WILLIAMSON, E. G. and BORDIN, E. S., " The Evaluation of Vocational and Educational Counseling: A Critique of the Methodology of Experiments." *Educ. Psychol. Measmt.*, 1, 1941.
WINSTON, J. T., " An Intensive Evaluation of the Services of a High School Guidance Program." *University of Missouri Bulletin*, 57: August, 1956.
WORBOIS, G. M., " Effect of Guidance Program on Emotional Development." *J. Appl. Psychol.*, 31, 1947.*U*
Year Book of Education. London: Evans Bros., 1955.

ENVIRONMENTAL CONDITIONS AFFECTING INTELLIGENCE

by R. Lynn
(Exeter University)

1. Introduction

It is now agreed by authorities such as Burt and Vernon that in western society inheritance is the major factor determining differences in what teachers understand by intelligence. This has been made clear by the two previous articles in the present series. Nevertheless, environment counts for a great deal and since education at home and at school acts through the environment, it is important to examine what we know about environmental influences. Relatively few studies have attempted to examine this question at all closely, partly because of the difficulty of isolating the suspected variables. However, there is enough evidence to suggest that there are at least five broad environmental factors of importance. This paper presents the evidence for these factors and discusses the mechanisms by which they work.

2. Schooling and Intellectual Stimulation

Perhaps the first place we should look for an environmental variable affecting intelligence is the school. However, since in western culture virtually all children go to school it is extremely difficult to determine the effect of this factor. We may plausibly assert that schooling increases the functional or operational intelligence and that universal compulsory education has certainly made the adult populations of western countries more able to think and act effectively in complex situations. The differences disclosed when scores on non-verbal tests achieved by technically advanced and educated groups are compared with those coming from underdeveloped countries are probably to be ascribed, at least in part, to differences in educational stimulation. Nevertheless, it must be admitted that it is difficult to show this in any conclusive way, since in the first place the so-called ' culture free ' tests may well be biased in favour of western populations; and secondly, the degree of intellectual attainment of a community is fairly certainly the result of the type of home conditions and the material standard of living as well as of education. Even within our own society differences between groups are difficult to interpret. For example, there is a well known study of Gordon showing that children brought up on canal boats with interrupted and intermittent schooling scored progressively lower on intelligence tests as they got older. This has frequently been interpreted as demonstrating the adverse effects of an irregular schooling on measured intelligence. However, it is now known from McLaren's study that the association of socio-economic status with intelligence gets progressively greater as the child grows older throughout the population, whether the child's schooling is irregular or not. It seems, therefore, that Gordon's findings could be due to any of the factors depressing intelligence test scores which are associated with the underprivileged home.

A field in which the effects of schooling have been effectively studied is nursery school education. Here a large accumulation of evidence indicates that, for children enjoying an ordinary family life, nursery school attendance has no effect on measured intelligence. Stoddard *et al.*, summarise nine studies of which eight report that there is no tendency for intelligence to be increased by nursery school attendance. Those studies such as Wellman's which appear to find increases in intelligence with nursery schooling can probably be explained in other ways. In this, 652 children attended pre-school in the U.S.A. for one year. The mean initial I.Q. was 117; and at the end of the year at pre-school the mean I.Q. was 124. 29 children attended the pre-school for 4 years; their mean initial I.Q. was 116, their mean final I.Q. 127. These findings may appear to suggest that nursery school attendance does something to raise the child's I.Q. as measured by tests. It should be observed, however, that these were children of high socio-economic status and since the association of I.Q. and socio-economic status probably increased with the age of the child, rises in I.Q. would be expected to occur even without the nursery school training.

On the other hand, gains both in measured intelligence and in operational ability may be present in special circumstances, such as where the intellectual stimulation is intense, or where, the home environment having been exceptionally poor in intellectual stimulation, the school makes up for it.

Certain researches support this hypothesis. McCandlass studied two groups of 6 four-year-old children with I.Q.'s over 124 who were matched for intelligence. The experimental group was given intense intellectual stimulation in which the children were given a great deal of personal attention and made to plan activities for themselves. After 6 months the mean I.Q. of these children had risen above that of the control group. It is likely, however, that these experiences were rather similar to test situations, and the gain in I.Q. might be regarded as a practice effect rather than as a genuine rise in real intellectual ability. It is also somewhat doubtful, bearing in mind the evidence from studies of the inconstancy of the I.Q., whether this gain would be maintained for long.

The most convincing demonstration of the stimulating effects of education comes from studies of orphanage children. In a study by Skeels, Updegraff, Wellman and Williams an entire orphanage of pre-school children was divided into two groups, matched for age, I.Q., sex, nutritional status, and length of residence in the orphanage. The age range was 1½-5½ years, and the mean I.Q. of the two groups about 80.

The control groups experienced the normal life of the orphanage and had little contact with adults. The experimental group lived exactly the same life, except that it spent the day in pre-school, in which there was contact with adults through teaching and play. Skeels and his collaborators present a number of results, of which a typical finding is the advance of 9 I.Q. points shown by the experimental group after nearly a year's (309 days) schooling (Table 1).

Table 1

	Number of Children	Binet I.Q. at Entrance	Age in months	Merrill Palmer I.Q. after 309 days	Standard Deviation
Pre-school . .	23	79	45	93	13
Control	23	79	44	84	12

In a similar study made by Dawe two groups of 11 orphanage children of pre-school age were matched and the experimental group was given intensive training, consisting of looking at books and pictures, listening to poems and stories and going on excursions with adults. Each child received an average of 50 hours of such special training. Table 2 shows the I.Q.'s of the two groups before and after training:

Table 2

	Initial S.B. I.Q.	S.B. I.Q. after Training
Experimental ..	81	95
Control ..	82	80

The differences in intelligence after the training of the experimental group is significant.

The evidence that among normal children nursery school experience does little to raise the I.Q. should probably not be interpreted to show that intellectual stimulation has no effect on real ability to learn but rather that there is not sufficient difference in intellectual stimulation for most children from adequate homes between staying with their mothers and going to nursery school. Where, however, the home is a poor one or where the child comes from an unstimulating orphanage environment, and if the difference in intellectual stimulation is large, considerable gains in measured intelligence may appear. There is no evidence, however, to show that these gains affect the child's learning ability or that they are permanent.

It is, of course, well known from the work of Vernon and others that direct coaching in intelligence tests may raise the I.Q. some 9 points; but this rise is regarded as spurious in the sense that the coaching does not carry over into other fields than those of test performance and even there does not last. Apart from this, it should not be too difficult to determine how far the best schooling can develop a child's effective ability to learn in comparison with the worst schooling. Two recent studies have some bearing on this question. Pidgeon, in this journal, brings together evidence which suggests that factors in the school may affect not merely measured attainments but measured intelligence as well. And the present writer reports a highly significant tendency for large schools to enhance attainment in the G.C.E. examinations.

3. Intellectual Quality of the Home

There is a considerable accumulation of evidence to show that there is a positive association between a child's score in intelligence tests and his performance in school and the intellectual and socio-economic quality of his home. Data of this sort, however, may be easily explained by the inheritance of ability and does not show unequivocally that the intellectual quality of the home is an environmental variable affecting the child's intelligence. For direct evidence isolating this variable we have to turn to studies of children brought up in foster homes of different intellectual status. If there is any association between a child's intelligence and that of his foster parents, this association cannot be due to heredity and must be ascribed to the effects of the environment. There are several studies showing that such an association exists.

In Leahy's study an investigation was made of the association between children's intelligence and the intelligence of their foster parents. There were 194 adopted children, all of whom were placed before they were six months old; at the time of testing there was an age range of 5-14 years. The mean I.Q., assessed by the 1916 Stanford Revision of the Binet, was 111. This group of children was matched with a control group of children living with their own parents.

The foster parents and the real parents were given the Otis Self-Administering Test and the Stanford-Binet Vocabulary Test. The intelligence of the parents was correlated with that of the children and the results are shown in Table 3.

Table 3

CHILD'S I.Q. CORRELATED WITH OTHER FACTORS

Correlated Factor	Adopted Children	Control Children
Father's Otis Score	.19	.51
Mother's Otis Score	.24	.51
Midparent Otis Score	.21	.60
Father's S.B. Vocabulary . .	.26	.47
Mother's S.B. Vocabulary . .	.24	.49
Midparent S.B. Vocabulary . .	.29	.56

The significant though relatively small correlations between the measured intelligence of the children and that of their foster parents suggests that the intellectual status of the parents is an environmental factor affecting the child's mental development. If the intellectual standard of the foster parents had no effect on the intelligence of the child, there should be no correlation between them. There are, however, two factors which may detract from the value of this evidence. The correlations may be spuriously high owing to the factor of selective placement, by which placement officers tend to put children of ' good ' parents into ' good ' foster homes. But probably this factor would not account entirely for the positive correlations. Secondly, intelligent foster parents may tend to have intelligent foster children because they possess some character quality rather than because they are intelligent as such. It might be that they set more store by intellectual inquiry and success and in some way pass this enthusiasm over to the child, so that the child works harder at educational tasks and strives to develop his own effective ability.

A study by Burks attempted to control the factor of selective placement in a study of 204 children adopted before the age of one year, at which age it is impossible to assess intelligence. The mean I.Q. of these adopted children was 107, and the correlation between their I.Q. and that of their adopting parents was +.20. This correlation is significantly lower than the correlation of +.50 between the I.Q.'s of children and parents in a control group of children brought up by their natural parents. Even though it is not entirely certain that selective placement has been quite eliminated in this study, it is unlikely that the whole of the difference in the correlations can be explained by it.

The other possibility, that it is the personality qualities of intelligent parents rather than intelligence as such, that affect the growth of the child's intellectual skills, is suggested by a study of Freeman, Holzinger and Mitchell. In this research 401 children were placed with adopting parents at an average age of 4.2 years. The investigators then related the intelligence scores of the adopted children to various indices of the prosperity, education and occupation of the adopting parents, and a correlation of +.48 was obtained. It is true that these children were adopted quite late when there are greater changes of selective placement, but this factor is to some extent controlled by the finding that the correlation was equally high in a group of 156 children placed before the age of two years. This correlation of .48 is markedly higher than those obtained by Burks and Leahy and this difference must be due to the different assessments of the adopting parents. It looks, therefore, as if the intelligence of the parents is a less important factor than the extent to which they attain middle class ideals; the higher correlation in the Freeman study suggests that the inculcation of middle class values may do more to develop the child's effective ability than purely intellectual stimulation as such. These values include such character qualities as persistence and ambition which undoubtedly motivate learning and may well affect the acquisition of skills required to score well in an intelligence test.

4. Contact with adults

Apart from the direct intellectual stimulation which adults may give either by teaching or by more informal day-to-day contacts, there is evidence that children who live in home conditions where there is a large amount of contact between mother and child do better on tests of intelligence and educational attainment. Here the influence is not so much the intellectual quality of the stimulation, as the quantity and quality of the attention and contact the mother gives, which directly affects the child's emotional and social growth, his feeling of security, and his willingness to put forth an intellectual effort.

That this is a factor is shown in an investigation by Skodak of the homes of 103 children brought up by foster parents. He assessed the characteristics of each home objectively by the use of a 'Home practice Inventory'. This inventory comprised ratings on a five point scale of such items as play equipment, social contacts, behaviour control, parents' cultural interests, the participation of the parents in the children's play, the frequency of play with other children, and the physical equipment of the home. This measure of the foster home environment correlated +.49 with the children's I.Q., assessed at the later pre-school age. The relationship of home inventory score to the child's I.Q. is shown in Table 4.

Table 4

HOME INVENTORY SCORE

	40+ to 49	50 to 59	60 to 69	70 to 79	80 to 89	90 to 99	100 to 109	110 to 119
Number ..	1	3	4	23	27	21	20	4
Mean I.Q.	87	90	102	109	113	113	121	127

This correlation of +.49 is somewhat higher than that found between the intelligence scores of foster parents and their foster children's I.Q.'s. This difference adds force to the suggestion that the emotional relationship and the amount of contact between parents and child may be a more important factor influencing the child's measured ability than parental intellectual stimulation as such.

A confirmatory study is that of Van Alstyne. Various aspects of the home were related to intelligence (Kuhlmann-Binet) of 75 three-year-old children, and it was found that the children's intelligence showed a higher correlation with the amount of parental contact than with the intellectual stimulation which comes from reading. The correlations obtained are shown in Table 5.

Table 5

Child's Mental Age by:	Coefficient of Correlation
1. "Opportunity for use of constructive materials"	+.50
2. Number of hours adults spend daily with child	+.32
3. Number of playmates in home	+.16
4. Number of hours father reads to child	+.06

This study shows clearly that the amount of adult contact is related to the test score of the child. But the study does not isolate this variable to the exclusion of others, and it might be that the parents who spend most time with their children are also more intelligent. If this were so, the superior intelligence of their children might be due to any number of causes. On the other hand, it may be that the interest parents take in their children is a personality trait not directly related to ability as such but a separate influence affecting ability.

Birth order and intelligence

The finding that eldest, youngest and only children are better achievers and more intelligent than intermediate children is indirect evidence supporting the hypothesis that the amount of parental or maternal contact is a variable affecting the child's intelligence. Studies of family position and intellectual attainment are of two sorts: (i) studies of eminent men, where more eldest sons have been found eminent than would be expected by chance and where more youngest sons have been found eminent than intermediate sons (Galton; Yoder; Ellis). (ii) Studies of children; Godfrey Thompson, in his Scottish Survey of 1947, found that the superiority of the only, eldest and youngest children was revealed clearly in the test scores. The mean scores for 3, 4 and 5-child families are shown in Table 6.

Table 6

	3-Child Families			4-Child Families				5-Child Families				
	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{3}{3}$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{3}{4}$	$\frac{4}{4}$	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{3}{5}$	$\frac{4}{5}$	$\frac{4}{5}$
Mean Score	38.85	37.65	38.43	36.15	34.63	34.40	36.31	33.14	31.27	32.01	32.67	33.64

Similar figures appeared for children from 6, 7 and 9 child families, although the findings are less reliable in families of these sizes. These differences can hardly be ascribed to an hereditary factor, since whether a child is an intermediate or youngest is a matter decided *after* his birth. Nor is there any obvious factor such as the age of the mother which could account for the superiority both of the eldest and the youngest. The explanation of this difference seems to lie in some environmental factor and this may well be the amount of contact the child has with adults; intermediate children would tend to get less attention from their parents than eldest and youngest. This hypothesis is supported by the finding that twins were less intelligent by about 5 I.Q. points than singletons. The survey provided evidence that this difference could not be due to family size or socio-economic level; it could be attributed to the greater attention which mothers are able to give single children.

A large-scale study of the intelligence of French children by Tabah and Sutter revealed a further tendency which supports the hypothesis that there is a relationship between maternal contact and children's intelligence. When two-child families were considered, it was found that the mean intelligence score of children was higher when there was a longer interval between the births. When there was a long interval, the children scored almost as highly as only children. This finding occurred within each of the five occupational categories into which the sample was divided, and so cannot plausibly be attributed to social class differences. This study confirms the Scottish finding that twins are less intelligent than singletons, even when socio-economic factors are controlled.

It is not at all certain how the influence of the amount of attention a child gets from adults should be explained. It might work simply through the greater amount of intellectual stimulation which a large amount of attention will afford. On the other hand, it may be that maternal attention has its effect primarily on the child's personality and only secondarily on his intelligence. This interpretation is supported by the evidence of maternal deprivation which will be reviewed in the next section.

5. Maternal Deprivation and Emotional Security

Many writers have regarded emotional security during the first five years of the child's life as an important influence affecting the child's intellectual development. The emotional security of a child depends on a constant and uninterrupted relationship between child and mother (or mother substitute) such as is more or less generally found in normal families but not in the older type of institution for orphans. A number of studies of the effects of maternal deprivation have been reviewed by Bowlby, who concludes "When deprived of maternal care, the child's development is almost always retarded—physically, intellectually and socially" (p. 13).

It is generally considered that the intellectual retardation of maternally deprived children is a secondary result of a more deep-rooted personality disturbance. Children separated from their mothers tend to become first extremely anxious for affection and attention, and then listless, apathetic, and emotionally withdrawn. The intellectual retardation is held to be a consequence of these emotional reactions.

It is not easy to find evidence which supports this theory to the exclusion of other explanations, although there is undoubtedly a large body of evidence which shows that children brought up in orphanages and thereby maternally deprived do tend to be below average in intellectual and scholastic attainment. This evidence is obviously open to several interpretations, of which perhaps the most plausible are: (1) children in orphanages are born from unintelligent parents and inherit their low intelligence. There is, in fact, evidence that children in orphanages do tend to have parents of low intelligence (e.g. Stoddard); (2) children in orphanages do not have so much contact with adults as children in families and consequently do not have the same opportunity to learn; (3) children in orphanages lack mother love and consequently develop inadequate personalities which adversely affect their intellectual development.

A series of studies by Spitz and his collaborators has been designed to show the ill-effects of maternal deprivation on the development (both intellectual and emotional) of young children. The early development of four groups of children in different environments was studied, namely: (*a*) a Foundling Home, an institution for children whose mothers were unable to care for them outside the institution. There were seven or eight children to one nurse, and the children were kept in solitary confinement for reasons of hygiene and had no contact with adults most of the day; (*b*) a Nursery, a penal institution for delinquent girls. The girls who gave birth to a child while they were in the institution were allowed to care for the child until it was one year old and there was considerable contact between mother and child; (*c*) Normal professional parents; (*d*) Peasant families.

The children were tested during the first four months of life and at the end of the first year by the Hetzer-Wolf baby tests. These tests are concerned with general physical and mental development, including intelligence, and give a Development Quotient similar to the I.Q. Spitz found that the maternally deprived (Foundling Home) children showed a marked fall in D.Q. at the end of the year and were greatly retarded in comparison with the other groups. These results are shown in Table 7.

Table 7

Environment	No. of cases	Development Quotients	
		Average 1-4 months	Average 9-12 months
Foundling Home	61	124	72
Professional ..	23	133	131
Peasant	11	107	108
Nursery	69	101	105

This work has been criticised by Pinneau on a number of grounds. In the first place, Pinneau points out that three studies have shown that the Hetzer Wolf tests are poorly standardised, in such a way that they are too easy in the early months and too hard in the later months of the first year. This faulty standardisation would have the result that any group of children would show a fall in D.Q. during their first year of life. However, Spitz's three control groups do not show such a fall and the difference remains to be explained. Pinneau further points out that the foundling home children were with their mothers during the first six months of life, and that almost all the drop in D.Q. occurred before the children were six months old, i.e. before they had been separated from their mothers. If this is so, it would be wrong to attribute the decline in D.Q. to maternal deprivation. Spitz does not reply very adequately to this criticism and he gives rather little information about the mothers of the Foundling Home children. It is also not clear whether Spitz has made longitudinal studies of the same children or whether the samples comprise different combinations of children at different ages.

Goldfarb attempted a study which would isolate the factor of maternal care in childhood, and control both hereditary influences and the intellectual stimulation that comes from contact with adults. In his investigation 30 children were examined, of whom 15 had been transferred direct from their mothers to foster homes within the first nine months of life, and 15 had been placed in an institution from about 6 months of age to 3½ years, and then transferred to foster homes. The institution had high standards of physical hygiene but provided very little adult-child contact. "Their only contacts with adults occurred during the few hurried moments when they were dressed, changed or fed by nurses." The mothers of the institution children were slightly superior to those of the foster home children, as assessed by educational, occupational and mental status. Differences between the two groups of children can therefore be ascribed to their different circumstances in infancy.

Goldfarb tested the two groups of children when their ages ranged from 10 to 14 years, and found that the children brought up in the institution for their first three years were retarded in comparison with the foster home children. Table 8 shows these differences, all of which are significant at the 1 per cent. level.

Table 8

Function Tested	Institution Children	Foster Children
Intelligence (Wechsler)	72·4	95·4
Ability to conceptualise:		
Weigh	2·4	6.8
Vigotsky	0.5	4·7
Reading	5·1	6.8
Arithmetic	4·7	6.7

These findings cannot very plausibly be explained by the intellectual stimulation theory. It might be supposed that the foster home children simply retained the lead which individual attention for two or three years in infancy would give them. But, considering the acknowledged inconstancy of the I.Q. between the ages of three and ten, it is doubtful whether this could account for the differences in the two groups at the ages of ten to fourteen. It seems most plausible to attribute Goldfarb's findings to the effects of deprivation on the personality, rather than to any lack of intellectual stimulation. As Bowlby points out, maternal deprivation tends to impair character development and produce a personality lacking in powers of concentration and persistence, and this might well reduce the level of operational or functional intelligence.

6. Attitude of Parents to Intellectual Attainment

It seems likely that parents who value intellectual and cultural achievements will tend to have intelligent children. It is parents of this sort who constitute what is known as ‘ a good home ’. This variable is not easy to isolate, however, since these parents are generally of higher than average ability themselves and the high ability of their children might be attributed to hereditary factors or to direct teaching. Yet there is some evidence that the parents’ attitude to intellectual values is a further factor affecting the child’s ability as shown in tests of all intellectual kinds.

The study of Kent and Davis has the great merit that it demonstrates a relationship between parental attitudes to educational success and the child’s measured intelligence and reading attainment, and at the same time attempts to show that the intelligence of the parents is not related to their attitude to success, thereby ruling out the factors of heredity and direct learning.

In this study, I.Q.’s and reading ages of 118 normal eight-year-old children were tested, and the mothers of these children were interviewed in their homes. On the basis of informal conversation, Kent classified the mothers into four groups according to the type of discipline and relationship which they imposed on the child. These four groups were as follows: (*a*) *Demanding* parents set high standards especially in school work, and rewarded ungenerously; (*b*) *Normal* parents were more tolerant and their expectations are related to the child’s needs, abilities and spontaneous interests, but they had positive standards; (*c*) *Overanxious* parents were ceaselessly anxious lest their child should fall short of what they expect, but their behaviour was inconsistent; (*d*) *Unconcerned* parents were indifferent to the child’s success and failure and their discipline was haphazard, inadequate and inconsistent.

Table 9 shows the number of parents in each group, the mean I.Q. of the children as measured by the revised Stanford Binet (Form M) and the Performance Scale of the WISC. and the mean reading age as assessed by Schonell’s Graded Reading Test. It is clear from the table that the parents who value intellectual standards and put pressure on their children produce the children with higher intelligence test scores, especially on the verbal side.

Table 9

	Normal	Demanding	Overanxious	Unconcerned
Number of children . .	41	38	30	9
Mean Binet I.Q.	109.9	124.2	107.3	97.0
Mean WISC I.Q.	110.4	113.4	100.7	103.0
Mean Schonell Reading Age	8.4	9.1	8.0	6.4

Many studies of this type are open to the objection that the similarities of parents and children could be explained directly by inheritance. This possibility is eliminated in the Kent and Davis study. The authors obtained information about the occupational status of the fathers and by dividing them into five groups according to the registrar-general's classification, showed that there was no significant tendency for the more ambitious parents to belong to the higher social classes.

Kent and Davis' study is supported by such evidence as we have about the discipline used by the parents of geniuses, for it is common to find that very talented people have been brought up in an atmosphere of great intellectual pressure. Before Beethoven was ten years old he was frequently made to play the violin through the night and when Thomas Arnold was three his birthday present was Smollet's *History of England* in twenty-four volumes. Cox has shown that the family histories of geniuses show that many of them have had ‘ demanding ’ parents in the sense used by Kent and Davis.

Socio-economic status and intelligence

The frequently found positive association of socio-economic status with intelligence and educational attainment (e.g. McLaren) might be attributed to the more favourable attitude of higher class parents to intellectual attainments. Although this association is often attributed to hereditary influences, the fact that it gets progressively larger with the age of the child suggests that environmental factors play some part. Indeed, two studies report that there is no association at all between socio-economic status and the measured ability of babies aged eighteen months (Furfy and Muchlenbein; Bayley), although the tests we have for this age are rather unreliable. But in so far as the effects of superior socio-economic status are environmental, they should probably be attributed either to the effect of the parents' attitude to intellectual success, or to the greater opportunities the children have for learning, or indeed to a combination of both.

III. Conclusion

It looks as if there are at least two mechanisms by which the environment affects a child's ability as shown in school and in the affairs of life. The first is the child's direct learning from the environment and copying from adults. The richer and more varied the intellectual stimulation, the greater the child's opportunity for learning and copying, and the more ‘ intelligent ’ he becomes. This mechanism would account for the tendency of a child's measured intelligence to be associated with intellectual stimulation, the intellectual quality of the home, and the amount of contact with adults.

The second mechanism is more complex and involves the whole personality. It seems as if certain environmental conditions in childhood can have a permanent effect on the personality, and this will in turn manifest itself in the way in which the child uses his ability. Only some process affecting the whole personality seems able to explain the effects on a child's measured intelligence and attainment, of maternal deprivation in infancy and the attitude of the parents to intellectual attainments.

In the state of present-day knowledge it is not at all clear how such a mechanism would work. In psychoanalytic terms, it might be explained by identification; if the parents have intellectual values, the child will incorporate these values into his superego in the same way as he incorporates standards of cleanliness and moral behaviour. When we are considering educational attainment, these acquired character qualities are especially important and the lack of them may account for many of the disappointing failures of clever working class children in England when they get to the grammar school. " Many," as Burt says, " seem to lack the spur of ambition." It may well be that not only educational attainment, but functional ability itself is enhanced by the right character qualities; and that the child acquires these from his parents through the process of identification.

In behaviourist terms these personality traits would be conceptualised as one or more acquired drives and intelligence itself as a system of habit strengths. This view of intelligence is consistent with the recent trend away from the view of intelligence as innate potential which intelligence tests measure inadequately and towards the conception of intelligence simply as the level of competence at which an individual is functioning at a particular time, as " the general all-round ability that an individual manifests in his daily life adjustments " (Vernon), or even as an acquired skill (Davis). The nature of the acquired drives contributing to the development of these habit strengths is more problematical. Terman cites such traits as persistence, application and the desire to excell, but are these all to be regarded as acquired drives? The problem of isolating and differentiating the many postulated acquired drives is in considerable confusion.

One of the most commonly postulated acquired drives, that of anxiety, probably has some effect on the development of intellectual skills. Although anxiety is commonly found to have a disorganising effect on complex learning in experimental situations, there is some evidence for Davis' theory that it acts as a motivating and socialising force in many real life learning situations (Burt; Lynn). Such evidence as there is about the effects of upbringing on the personality of the child supports this hypothesis. It is generally held that the child who has a great deal of maternal attention (sometimes called maternal over-protection) becomes nervous and anxious (e.g. Mayer Gross, Slater and Ross, Levy; Banister and Ravden). As we have seen maternal attention also tends to produce children who are ' intelligent ', and maternal deprivation children who are ' unintelligent'. Davis' theory would explain this as follows. The attention and love which mothers give their children make them concerned to obtain adult approval and educational success. This concern is what Davis calls socially necessary anxiety, and it acts as a motivating force making children put effort into their school work. As a result of this they learn more quickly the intellectual skills that secure a high intelligence quotient.

The question remains of what the schools can do to raise their children's intelligence and attainment levels. The evidence suggests that the basic personality patterns are laid down in the home at an age before the children arrive at school. The schools' task seems chiefly to arouse the potential of motivation which is present in the great majority of children. Traditionally, this has been done by the competitive system, by arousing the interests of the children, and by the personality of the teacher arousing inspiration. It must be admitted that at the present we have much speculation but very little precise knowledge of the relative effectiveness of these motivators and it is perhaps in this field that some of the most urgent research needs to be done.

References

1. BANISTER, H. and RAVDEN, M., " The problem child and his environment." *Brit. J. Psychol.*, 1944, 34, 60-65.
2. BAYLEY, N., *Mental Growth during the First Three Years: a developmental study of sixty-one children by repeated tests.* Genet. Psychol. Monogrs., 1933, 14, 1-92.
3. BOWLBY, J., *Maternal Care and Mental Health.* World Health Org., Monograph Series, No. 2, Geneva, 1951.
4. BURKS, B. S., " The relative influence of nature and motive on mental development: a comparative study of foster parent–foster child resemblance and true parent–true child resemblance." *Year B. Nat. Soc. Stud. Educ.*, 1928, 27 (1), 219-316.
5. BURT, C., *The Backward Child.* London: Univ. London Press, 1937.
6. BURT, C., *Intelligence and Fertility.* London: Eugenics Society and Hamish Hamilton, 1946.
7. BURT, C., Symposium on the selection of pupils for different types of secondary schools. *Brit. J. Educ. Psychol.*, 1947, 17, 57-71.
8. COX, C. M., *The Early Mental Traits of Three Hundred Geniuses.* London: O.U.P., 1926.
9. DAVIS, A., " Socialisation and Adolescent Personality." In *Readings in Social Psychology*, ed. Newcomb, T. M. and Hartley, E. L. New York: Holt, 1944.
10. DAVIS, D. R., " Nurture and Intelligence." *Bull. Camb. Inst. Educ.*, 1954, 1, 1-9.
11. DAWE, H. C., " A study of the effects of an educational programme upon language development and related mental functions in young children." Univ. Iowa: Unpublished *Ph.D. thesis*, 1940.
12. ELLIS, H., *A Study of British Genius.* Boston: Houghton Mifflin, 1926.
13. FREEMAN, F. N., HOLZINGER, K. J. and MITCHELL, B. C., " The influence of environment on the intelligence, school achievement and conduct of foster children." *Year B. Nat. Soc. Stud. Educ.*, 1928, 2 (1), 103-217.
14. FURFY, P. H. and MUCHLENBEIN, J., " The validity of infant intelligence tests." *J. Genet. Psychol.*, 1932, 40, 219-223.
15. GALTON, F., *English Men of Science: their nature and nurture.* London: Macmillan, 1874.
16. GOLDFARB, W., " Effects of early institutional care on adolescent personality." *J. Exp. Educ.*, 1943, 12, 106-129.
17. GORDON, H., " Mental and scholastic attainments among retarded children: an inquiry into the effects of schooling on the various tests." Educ. Pamphlets, Bd. Educ., London, No. 44, 1923.
18. JONES, H. E., " Environmental influences on mental development." In *Manual of Child Psychology*, ed. Carmichael, L. New York: Wiley, 1946.
19. KENT, N. and DAVIS, D. RUSSELL. " Discipline in the Home and Intellectual Development." *Brit. J. Med. Psychol.*, 1957, 30, 27-33.
20. LEAHY, A. M., *Nature-nurture and Intelligence.* Gent. Psychol. Monogr., 1935, 17, 236-308.
21. LEVY, D. M., *Maternal Overprotection.* New York: Columbia Univ. Press, 1943.
22. LYNN, R., " Personality factors in reading achievement." *Proc. Roy. Soc. Med.*, 1955, 48, 996-998.
23. LYNN, R., " The Relation of School Size to Educational Attainment. *Brit. Journ. Soc.* (in the press).
24. MAYER GROSS, W., SLATER, E. and ROTH, M. *Clinical Psychiatry.* London, 1954.
25. MCCANDLASS, B. R., " The effect of enriched educational experiences upon the growth of intelligence of very superior children." University of Iowa, 1938. Unpublished *M.A. thesis.*
26. MCLAREN, V. M., " Socio-economic status and reading ability—a study in Infant reading." In *Studies in Reading*, Vol. II. Scottish Council for Research in Edu., 1950, No. 34. London: Univ. London Press.
27. PEEL, E. A., *The Psychological Basis of Education.* London: Oliver and Boyd, 1956.
28. PIDGEON, D. A., " School differences in ability and attainment." *Ed. Res.*, Vol. I, No. 3.
29. PINNEAU, S., " The infantile disorders of hospitalism and anaclitic depression." *Psychol. Bull.*, 1955, 52, 429-462.
30. SKEELS, H. M., UPDEGRAFF, R., WELLMAN, B.L. and WILLIAMS, H. M., " A study of environmental stimulation: an orphanage preschool project." Univ. of Iowa, *Stud. Child Welfare*, 1939, 16, No. 1.
31. SKODAK, M., " Children in Foster Homes." Univ. Iowa, *Stud. Child Welfare*, 4, 1938.
32. SPITZ, R. A., " Hospitalism: a follow-up report." *Psychoanal. Stud. Child*, 1946, 2, 113-117. New York: International Univer. Press.
33. SPITZ, R. A., " Anaclitic depression." *Psychoanal. Stud. Child*, 1946, 2, 313-342. New York: International Univer. Press.
34. SPITZ, R. A., " Reply to Dr. Pinneau." *Psychol. Bull.*, 1955, 52, 453-459.
35. STODDARD, G. D., *et. al.*, " Environment and the I.Q." The 39th Year Book Nat. Soc. Stud. Educ. *Intelligence: its Nature and Nurture.* Dr. I. Bloomington: Public School Publishing Co., 1940.
36. STODDARD, G. D., *The Meaning of intelligence.* New York: Macmillan, 1943.
37. TABAH, L. and SUTTER, J., " Le niveau intellectuel des enfants d'une même famille." *Ann. Hum. Genet*, 1954, 19, 120-150.
38. TERMAN, L. M., *et. al.*, *Genetic Studies of Genius.* Stanford University Press, 1926.
39. THOMPSON, G., " Intelligence and Family Size. In *The Trend of Scottish intelligence.* London: Univ. London Press, 1949.
40. VAN ALSTYNE, D., *The Environment of Three Year Old Children.* Teach. Coll. Contrib. Educ., No. 366, 1929.
41. VERNON, P. E., " Recent Investigations of Intelligence and its Measurement." *Eugenics Review*, 1951, 125-137.
42. VERNON, P. E., " Symposium on the effects of coaching and practice in intelligence tests." *Brit. J. Educ. Psychol.*, 1954, 24, 57-63.
43. WELLMAN, B. L., *Iowa Studies on the effects of schooling.* 39th Year Book Nat. Soc. Stud. Educ., 1940.
44. YODER, A. H., " A study of the boyhood of great men." *Ped. sem.*, 1894, 3, 134-156.

SCHOOL TYPE DIFFERENCES IN ABILITY & ATTAINMENT

by D. A. Pidgeon

Introduction

THE organisation of schools in this country is such that at any given age it is possible for a child to receive his education in one of two or more different types of school.

At the secondary stage, apart from one or two small areas with comprehensive schools only, this variety of provision is deliberate, and differences in the abilities and attainments of children attending grammar and modern schools are to be expected. Interest is only aroused by the extent to which performances tend to be similar, that is, by the amount of overlap that exists between the schools. At the primary level, however, educational objectives are similar for all children, yet some will be attending Junior schools with an Infants department attached under the same Head, some will be in Junior schools taking children from 7 to 11 only, and others will still be receiving their education in unreorganised, all-age schools.

It is, therefore, a matter of concern to administrators, teachers and parents to know whether attendance at one of these different types of primary school exerts a greater or lesser influence upon the attainments of its pupils than attendance at any other. The purpose of this article is to bring together evidence that may throw light on this particular problem. Related questions, such as whether it is more or less beneficial to a child to be educated in a single-sex or mixed school, will be dealt with on a later occasion.

Although three types of primary school have been mentioned, the main concern will be a comparison between the Junior with, and the Junior without Infants schools. Any further comparison with All-age schools is really of academic interest only, since up to the age of eleven this type of school is not basically different from the Junior with Infants school and, in any case, it is now the accepted policy to replace them with one of the other types. There is, however, no similarly accepted policy with regard to the choice between Junior with, and Junior without Infants schools. Each Local Education Authority makes the decision, and apparently the main consideration is the size of school needed for the district in which it is to be situated: in the more densely populated areas the choice tends to be the Junior without Infants type of organisation; in the more sparsely populated areas, Junior with Infants.

An examination of the Ministry of Education's statistical returns reveals the tendency for the former type of school to be favoured in recent years. Table I shows the percentage of all children attending the three types of primary school in 1948 and in 1956.

Table I

PERCENTAGE OF ALL CHILDREN ATTENDING JUNIOR WITH INFANTS, JUNIOR WITHOUT INFANTS AND ALL-AGE SCHOOLS IN 1948 AND 1956

YEAR	JUNIOR WITH INFANTS	JUNIOR WITHOUT INFANTS	ALL-AGE
1948	34.64%	30.54%	34.82%
1956	42.86%	42.32%	14.82%
Change '48-'56	+8.22%	+11.78%	—20.00%

During the years 1948-56 the proportion of children attending All-age schools has decreased considerably—in accordance with declared policy, but although there has been a corresponding increase in both other types of school, relatively more children are now being educated in separate Junior schools. The precise reasons for this are obscure, but the possibility of it being related to density of population can be examined by looking at the Ministry's returns for Urban and Rural areas separately. These are given in Table II.

Table II

PERCENTAGE OF ALL CHILDREN ATTENDING JUNIOR WITH, AND JUNIOR WITHOUT INFANTS SCHOOLS IN URBAN AND RURAL AREAS FOR 1948 AND 1956

YEAR	URBAN		RURAL	
	J. with I.	J. without I.	J. with I.	J. without I.
1948 1956	46.8% 42.2%	53.2% 57.8%	86.6% 82.8%	13.4% 17.2%
Change '48-'56	—4.6%	+4.6%	—3.8%	+3.8%

Table II suggests that the increasing tendency for more children to be educated in separate Junior schools is almost as marked in rural as in urban areas. Hence density of population is probably not the only cause for preferring one type of school to the other.

There is no doubt, however, that for whatever reason, an increasing number of children today aged 7-11 are receiving their primary education in separate Junior schools. This is an interesting trend, since from the purely educational viewpoint cogent reasons have been advanced for preferring the type of organisation which provides unbroken tuition under one Head Teacher from entry at 5 to the beginning of secondary education at 11.

It can be argued that for some children at least the necessity to change to a different, invariably larger, school at the still tender age of 7 can be an emotional experience that might have an adverse effect on their attitude to school and school work. Morris, in dealing with this problem of transition, suggests that for many children the transfer from Infant to Junior schools comes too early, and she quotes a report of the National Association for Mental Health:

> " The effects of too-early transition are seen in an increase both of emotional troubles, such as nail-biting, enuresis, fear of school, spitefulness; and of learning difficulties, for example, a standstill in reading, free-writing and number. These bad effects are, of course, the more marked where the change of atmosphere is the greater, when, for example, a somewhat sentimental infant school is followed by a harsh junior school."

More direct evidence of the difficulties likely to be encountered in making a transition to separate junior schools is given by Morris. From her survey of reading in Kent schools she found, using the stage reached in progressing through reading primers as a criterion, that " at the normal time of transfer, approximately half of the children had reached a stage of development from which they could make progress themselves if given encouragement in a stimulating environment. The rest, however, still needed the kind of teaching associated with the infants school." But she also found that whereas 43 per cent. of teachers taking the transition classes in Junior with Infants schools had had specific training in infant methods, the corresponding proportion in Junior without Infants schools was only 17 per cent. Similarly, schools catering for children aged five to eleven had a higher proportion of first-year junior teachers who had had experience in an infants school. Also, as Morris points out, there is a far greater opportunity for direct liaison between top infant and first-year junior teachers if they are both within the same school building. Indeed, she reports that the Head Teachers of the separate junior schools she interviewed all felt that for real continuity of teaching methods the Junior with Infants type of organisation was much more satisfactory.

No one questions the importance attached to the teaching of " slow learners," but to ensure the steady progress of such children it is not only necessary to maintain continuity of method but appropriate teaching materials should also be available to them for as long as they are required. Separate junior schools may be at a disadvantage in this respect, since not all will find it possible to provide the teaching apparatus necessary for what may be only a small proportion of their children.

Finally, there are always a few emotionally immature children who could with advantage remain longer, possibly for a term or even a year, in an infants department. The retention of such children is administratively much easier if no change is made to a separate school.

It can, of course, be maintained that the mere fact of promotion to junior status will act as a stimulus to many children and encourage them to take a greater interest in school work. But this incentive would apply equally to both types of organisation, and it is doubtful whether any additional advantage that might result from achieving higher status in a separate school would outweigh the many disadvantages that appear to result from making such a transition.

Research evidence

If there is substance in the foregoing discussion, an examination of the relative performances of children in the different types of primary school should reveal the superiority of those lucky enough to receive their education in Junior with Infants schools. The necessary evidence, however, can only be obtained where fairly large and representative samples of children attending different types of school have been examined with similar types of test and the results analysed so that the differences can be compared. Unfortunately, the number of investigations that have been carried out, involving the surveying of the ability and attainment of primary school children, is comparatively small and most of those that have been published do not contain a break down of the results that permits a comparison of the kind required.

The Ministry of Education, however, has published the results of its 1956 Reading Survey of eleven-year-old children with separate means for the different types of primary school. The relevant figures are reproduced in Table III.

Table III

MEAN SCORES ON THE MINISTRY'S WATTS-VERNON READING TEST BY TYPE OF SCHOOL WITHIN TYPE OF AREA

N=1,374 Children aged 11+

TYPE OF AREA	TYPE OF SCHOOL		
	J. with I.	J. without I.	All-age
Urban ..	13.6	13.8	12.8
Rural ..	12.1	13.0	12.0
All areas ..	13.2	13.7	12.6

The Ministry report adds, " Differences of the same order between types of school . . . were found in the other two surveys " (1948 and 1952). None of the differences in Table III is, in fact, statistically significant, but these results certainly do not confirm the expected superiority of Junior with Infants schools.

Further evidence can be obtained from a reading survey of all children in the 11+ age group carried out, with the assistance of the Foundation, by a County Education Authority. In this survey the Foundation's Sentence Reading Test 1—a test of reading comprehension—was given and estimates were made of the proportion of children in the area defined as " illiterate." The purely arbitrary definition of illiteracy adopted for these eleven-year-old children was a score on the reading test of 7 or under, since the median score on the same test for seven-year-old children was between 7 and 8 points. Table IV gives the proportion of " illiterates " among children attending different types of primary school in urban and rural areas.

Table IV

PERCENTAGE OF " ILLITERATES " BY TYPE OF PRIMARY SCHOOL AND TYPE OF AREA

N=3,215 Children aged 11+

TYPE OF AREA	TYPE OF SCHOOL		
	J. with I.	J. without I.	All-age
Urban ..	5·1	2.7	3·4
Rural ..	4·7	1.6	3·6
All areas ..	4·83	2.83	3·55

The Junior without Infants schools show a significantly smaller percentage of " illiterates " than the Junior with Infants schools. Differences between Urban and Rural areas are not significant. In this same investigation, when a higher standard of reading was taken to define " backward " readers, similar differences were found. It would appear, therefore, as if the Ministry's findings are confirmed. In reading, at least, the performance of children in separate Junior schools appears to be superior, and in the face of the reasoning given earlier this is certainly unexpected. As in all research work, test were similarly associated with schools in good districts, of large size and with superior buildings. They were also associated with a Junior without Infants type of organisation, but to a lesser degree of statistical significance.

These results agree with the other evidence cited earlier. Morris, however, carried out further analyses of her data aimed to eliminate the effect of those variables which cannot readily be altered, namely, urban/rural location, socio-economic status and size of school. The association between good reading and the Junior without Infants organisation was then shown *not* to be statistically significant.

In addition to using the scores on a reading test alone as a criterion of performance, Morris also used " reading adjusted for intelligence," the scores on non-verbal tests being employed for the adjustment. Her analyses showed that when allowance was thus made for differences in " intelligence," there was, nevertheless, a marked and highly significant association between higher reading ability and the Junior without Infants type of organisation. This remained even after eliminating the effects of urban/rural location, socio-economic status and school size. Assuming that the adjustment process was valid, this would indicate that Junior without Infant schools were indeed producing better readers.

So far, all the evidence considered has come from the testing of eleven-year-old children, or, in Morris' case, all children aged 7 to 11. However, in Morris' very thorough investigation a further series of analyses was carried out on a sample of first-year junior children only. In this sample, as the children had been juniors for only six months, it is hardly likely that good attainment could be attributed to the influence of the junior school, but the results again showed a highly significant association between good reading (and good non-verbal scores) on the one hand, and Junior without Infants schools, superior buildings, high socio-economic status and large school size on the other. Moreover, when the effect of the last two variables (and urban/rural location) were partialled out from the analysis a significant association remained between good reading and the Junior without Infants type of organisation. This last result is certainly surprising since, as Morris puts it, " Unless progress can be attributed to a change of school, the fact of being educated in a separate infants school is perhaps more important for reading development than later attendance at a school for juniors only." And she goes on, " These results would, therefore, appear to refute the common supposition that a Junior with Infants type of organisation is better for ensuring continuity of development."

From the evidence reviewed so far it seems fairly clear that children attending separate Junior schools produce better reading test results than those attending Junior with Infants schools or unreorganised All-age schools. It also seems fairly clear that the separate Junior schools on the whole are larger, have superior buildings and are to be found in neighbourhoods of high socio-economic status. Whether the higher reading scores are due entirely to this latter fact or whether the separate Infant and Junior organisation affords additional benefits on the children being educated in them is not so clear.

Evidence from the Foundation's National Survey

In 1955, as a part of a National Survey of Ability and Attainment, fully representative samples of children in three age groups were given tests of reading, arithmetic and non-verbal ability. Of interest to the present inquiry are the results obtained from the 7+

and 11 + age groups. The former were tested during October, that is, about six weeks after they had become "juniors." The latter were tested in May, that is, during their last term as "juniors."

The 7 + age group was given a non-verbal picture test, a test of reading and a test of mechanical arithmetic which consisted of four separate sub-tests of 20 items each in addition, subtraction, multiplication and division. The 60-item reading comprehension test was the same as that given to children in the older age groups of the survey and although it proved to be highly successful with these latter groups it was found not to give effective discrimination with the younger children and hence, unfortunately, had to be omitted from the subsequent analyses.

The mean standardised scores obtained by the 7 + children for the non-verbal and arithmetic tests are given in Table VI.

Table VI

MEAN STANDARDISED SCORES ON THE NON-VERBAL AND ARITHMETIC TESTS FOR CHILDREN ATTENDING DIFFERENT TYPES OF PRIMARY SCHOOL

N=2,417 Children aged 7 +

TEST	TYPE OF SCHOOL		
	J. with I.	J. without I.	All-age
Non-verbal	98.59	101.66	97.17
Arithmetic	101.28	99.84	99.71

Although the differences for the most part just fall short of statistical significance, the results are nevertheless extremely interesting, especially when it is remembered that these children had only just become "juniors." The children in the separate junior schools, although giving a better performance on the non-verbal test, did not do as well as the combined Junior with Infants schools in arithmetic. Care should be taken not to place too much weight on this result in view of its statistical non-significance, but it does apparently provide some counter-evidence, at least in arithmetic, to Dr. Morris' conclusion that for reading education in a separate Infants school is advantageous.

The 11 + age group completed a reading comprehension test and two tests of arithmetic. The first of the arithmetic tests (Arithmetic 1) was a 25-item "mechanical" test ranging from simple addition to multiplication and division of fractions and decimals. It was given without time limits. The second arithmetic test (Arithmetic 2) consisted of 100 items, 40 mechanical and 60 problem, administered with a time limit of 50 minutes. This test was in fact similar to the usual type of arithmetic test used by many Local Education Authorities as part of their 11 + allocation procedures. The children in this age group also completed a 64-item non-verbal test. The mean scores for this age group are given in Table VII.

Table VII

MEAN SCORES ON FOUR TESTS FOR CHILDREN ATTENDING DIFFERENT TYPES OF PRIMARY SCHOOL

N=3,029 Children aged 11+

TEST	TYPE OF SCHOOL		
	J. with I.	J. without I.	All-age
Reading ..	99.65	100.42	99.33
Arith. 1 ..	98.39	101.56	99.80
Arith. 2 ..	98.08	101.97	99.36
N/V ..	100.42	101.09	97.99

Again, not all the differences between the Junior with and without Infants schools reach statistical significance, although that for Arithmetic 2 does. The trend is, however, consistent, and apparently confirms the previous findings for children of this age.

In order to discover whether the higher mean scores for the Junior without Infants schools are due to the influence of the type of organisation or to the fact that this type of school contains " better " children, it is necessary to carry out further analyses. Unfortunately, no data on the socio-economic status of the catchment areas for the schools concerned were collected in the National Survey, but the size of each school, or rather, the size of the age group in each school, was known and, as both Kemp and Morris have shown, this variable is highly correlated with the type of area in which the schools are situated. An examination of the data showed that the average number of pupils in the age group concerned attending both Junior with Infants and All-age schools was just under 20 compared with nearly 80 for the Junior without Infants schools.

Accordingly, the mean test scores on Reading and Arithmetic Test 2 for different sized age groups in each type of school were examined. No association between test score and size was found in the All-age schools, but in Junior without Infants schools the association was highly significant for both tests, whilst in the Junior with Infants it was significant for arithmetic and nearly so for reading.

With arithmetic, when the effect of the size of the age group was eliminated by using a regression analysis, no difference between the mean scores for Junior with and without Infants schools remained—although the difference had been significant previously. The mean standardised scores on Arithmetic Test 2 for different sizes of age group are given in Table VIII.

Table VIII

MEAN STANDARDISED SCORES ON ARITHMETIC TEST 2 FOR DIFFERENT SIZES OF AGE GROUP IN BOTH JUNIOR WITH INFANTS AND JUNIOR WITHOUT INFANTS SCHOOLS

N=2,642 Children aged 11+

TYPE OF SCHOOL	SIZE OF AGE GROUP				
	20	40	50	60	80
J. with I.	96.0	97.8	98.7	99.6	101.4*
J. without I. ..	95.8*	97.6	98.5	99.4	101.2

*These values have been estimated from the data, since there was only one Junior with Infants school with more than 80 pupils in the age group, and no Junior without Infants schools with less than 30 pupils.

In the case of reading, however, eliminating the effect of the size of the age group reveals significant differences between the two types of school concerned. This time they are in favour of the Junior with Infants type of organisation. The relevant mean standardised scores are given in Table IX.

Table IX

MEAN STANDARDISED SCORES ON READING TEST FOR DIFFERENT SIZES OF AGE GROUP IN BOTH JUNIOR WITH INFANTS AND JUNIOR WITHOUT INFANTS SCHOOLS

N=2,642 Children aged 11+

TYPE OF SCHOOL	SIZE OF AGE GROUP				
	20	40	50	60	80
J. with I.	97·5	99·4	100.4	101.4	103.4*
J. without I. ..	93·7*	95·7	96.6	97.6	99.6

*These values have been estimated from the data, since there was only one Junior with Infants school with more than 80 pupils in the age group, and no Junior without Infants schools with less than 30 pupils.

The results set out in Tables VIII and IX are extremely interesting. Clearly, the size of the age group *per se* does not cause the difference in mean score; this must be the result of other variables with which size is associated.

It would seem, however, as if the superiority in performance of children attending Junior without Infants schools which is normally found is not due to any specific influence which this type of school may exert on its pupils. Indeed, the results of Table IX suggest that, other things being equal, the Junior with Infants type of organisation is to be preferred, at least so far as reading is concerned.

Summary and Conclusions

Although the trend since the war seems to be in favour of providing primary education in separate Infants and Junior schools, from the purely educational viewpoint a strong case can be made out for maintaining greater continuity in combined Junior with Infants schools. Evidence as to whether there are differences in the levels of attainment of children in the two types of organisation is meagre, but what there is appears at first sight to contradict expectation and to favour the separate Junior school. A closer examination of the problem suggests that the apparent superiority in performance of children attending Junior without Infants schools may be ascribed, not to any influence that this type of school might exert, but to the fact that separate schools for children aged 7 to 11 have been built for the most part within the last 20 or 30 years and are to be found in newer housing estates and better areas generally. As a result they probably contain a higher proportion of more able children.

In making a choice between the two types of organisation it is clearly of prime importance that due consideration should be given to the needs and interests of the children themselves. It could be maintained that administrative arguments for building separate schools should be waived in favour of a type of organisation that offers psychological and educational benefits, even if the final attainments of children attendi separate junior schools are demonstrably higher.

However, when appropriate allowances are made for the environmental fact associated with the type of area in which schools are built, the results, although in p conflicting, suggest that there is little difference in the attainments of children attendi the two types of organisation. There is even some indication that the performance children in combined Junior with Infants schools is superior. Although it would desirable to confirm this latter finding with evidence from further surveys, it would, nev theless, seem that the policy of building separate Infants and Junior schools needs to seriously questioned.

References

KEMP, L. C. D., " Environmental and other characteristics determining attainment in primary schools," *Brit. Educ. Psych.*, XXV, 67-77, 1955.

NATIONAL ASSOCIATION FOR MENTAL HEALTH, *Periods of Stress in the Primary School*, London, 1955.

MINISTRY OF EDUCATION, *Standards of Reading*, 1948-1956, H.M. Stationery Office, 1957.

MORRIS, J. M., *Reading in the Primary School*, Newnes, 1959.

BIBLIOGRAPHY

The curriculum of secondary schools, with particular reference to schools having courses biased towards technical education.

by F. Barr

IN "*Technical Education for Adolescents*" Dobinson stresses the difficulties of individuals trying to solve everyday teaching problems without the co-operation of their colleagues. Many of the following items have been selected to help those who would like to know how colleagues are attempting to solve problems which all teachers meet. The bibliography is not meant to provide ready-made solutions but to stimulate thought by disseminating ideas which have been developed and found useful by people in similar situations, as well as to indicate various sources of information for further reference.

The items of the bibliography have been arranged to include general works which concern the curriculum as a whole as well as works which refer to individual subjects, themselves grouped into general works with a technical or commercial bias, workshop arts and crafts, and commercial subjects. The list does not include all items concerning individual subjects or all subjects of the curriculum, and in particular there is no reference to the basic subjects of English and Mathematics which require separate bibliographies.

Many teachers have little spare time to seek information amongst the many present-day educational publications and the bibliography has been compiled with this problem in mind, for generally the articles are short in length and have been published in recent years apart from one or two classics which are still of great value. All should be readily available in local libraries of repute but if they are not, the librarians of teachers' Professional Organisations and of the local Institute of Education should be able to help. If local resources fail the Information Officer of the National Foundation for Educational Research can be consulted.

Among the sources of items in this bibliography are:

1. Journal of Education (J. of Ed.).
2. Vocational Aspects of Secondary and Further Education (Voc. Asp.).
3. Handbooks of the Incorporated Association of Assistant Masters (I.A.A.M.).
4. Her Majesty's Stationery Office Publications (H.M.S.O.).
5. British Association for Commercial and Industrial Education (B.A.C.I.E.).

The compiler is indebted to A. J. Peters for some of the following items.

GENERAL WORKS

Board of Education Pamphlet No. 114. H.M.S.O. London, 1928.

This pamphlet contains chapters on mathematics, science, engineering subjects, commercial work, art, domestic subjects and handicrafts. It assumes that whereas the early and middle period of the secondary school course is essentially an "Age of Romance," followed by an "Age of Precision," in A. N. Whitehead's terminology sixth form work is an "Age of Generalisation" vocationally biased whether in grammar or technical schools. It recommends links with the technical colleges and accordingly advises sound preliminary training in workshop practice, geometrical and machine drawing and science biased towards mechanical and electrical engineering. Mathematics and a foreign language are included in the scheme of work. It recommends that sixth form commercial courses should be a continuation of " general " education. Advice is given on the teaching of shorthand, typewriting and book-keeping. Courses in Domestic Science are also described.

" *The Second-Hand Curriculum.*" NEAL, L. F. **J. of E.** April, 1957. Vol. 89. No. 1053.

The author considers that many secondary school children receive an education which is based upon a second-hand copy of the grammar school curriculum, which is itself being subjected to critical assessment. An example is given of a rural secondary modern school which uses a course of rural studies as the core of the curriculum. The claim is made that by engaging in farming techniques pupils will appreciate the importance of the more academic studies.

The Secondary Technical School and its Curriculum. Nottinghamshire Education Committee: Development Sub-Committee Advisory Panel. Nottingham, 1949.

The report advocates a curriculum which shall be common to all pupils in the 11-13 year age-groups. Afterwards, some vocational bias is recommended but is to take not more than one-third of the timetable, the remaining two-thirds being devoted to " general " education. The type of vocational bias is to be determined by the pattern of local industry, thus it recommends courses biased towards engineering, building, mining, agriculture, commerce, domestic science and art. The study of foreign languages is recommended for those who can profit by it. Timetables and syllabuses are considered in some detail.

Secondary Education, with special reference to Grammar Schools and Technical High Schools (Spens Report). H.M.S.O. London, 1928.

The curriculum of the grammar school is considered thoroughly, in general composition and balance (p. 277) and by individual subjects (chapters iv-vi). Chapter viii includes a general consideration of technical schools and in particular considers the problems of the balance of " vocational " and " general " subjects. It gives advice on the curriculum of technical schools, modifying the basic pattern of that of the grammar school which is explained in detail. The report stresses the need of the technical school to develop its own traditions and courses of study to prevent it from becoming a pale copy of established grammar schools.

Education—intellectual, moral and physical. SPENCER H. WATTS. London, 1924 (first published 1861).

Spencer's main attack was directed against the dominance of classical subjects in the curriculum. The essays advocate an extreme form of " utilitarianism," although in his search for the knowledge which is of most use to men his interpretation of the word " use " is wide. This is an extremely stimulating series of essays about educational topics which are still very much discussed.

Science and Education. Collected Essays. HUXLEY, T. H. Vol. III. MacMillan. London.

Huxley was a vigorous protagonist of technical and scientific studies, but unlike Spencer he considered them to be not the necessary alternative but an integral part of a truly " liberal " education. Although old, some of the essays are good expositions of this point of view which is still a topic of argument today.

" *Staying on to Sixteen.*" Crofts, F. A. **B.A.C.I.E. Journal**, Vol. 10, No. 6. December, 1956.

The author considers that within our present industrial society, schools have a duty to provide for the full development of the individual, to enable him to fulfil his obligations to society and to provide for society's future development. These broad concepts are briefly analysed and expanded and their implications for the secondary school curriculum are studied with reference to industrial needs. The particular problems of staying on to sixteen, including those of taking external examinations and the organisation of courses suitable for this particular age-group, are examined briefly.

Technical Education for Adolescents. Dobinson, C. H. Harrap. London, 1951.

Besides general comments on industry and education and the historical development of apprenticeship, secondary technical and modern schools, the author has valuable comments concerning " Making Technical Education Liberal " in chapter VII. The comparisons with French methods are stimulating.

Reconstruction in the Secondary School. Earle, F. M. University of London Press. Bickley, 1944.

This is one of many books published during this period about the reconstruction of secondary education. Much of the work has special reference to Comprehensive Schools, but the chapters on the choice of curriculum with its sane plea to translate " high sounding generalities such as ' education for life ' into practical procedures " has important implications for secondary technical schools. This theme is further developed in the next two chapters on the construction of work programmes and education for employment.

A Review of Junior Technical Schools in England. Board of Education Pamphlet No. 111, H.M.S.O. London, 1937.

The general situation of Junior Technical Schools is surveyed as it existed at the time of publication. The curricula of these schools are considered on pages 17-25. The pamphlet states that although primarily vocational in purpose, curricula should not be confined to craft and technical subjects but should include those which pay attention to broader educational aims. Specimen curricula are given. Although the views expressed are not necessarily those of current educational thought some of its comments are still useful for teachers in these schools.

Curriculum and Examinations in Secondary Schools (Norwood Report). H.M.S.O. London, 1943.

The Report assesses the pre-war situation of secondary schools and makes recommendations for future developments. In particular it considers that the secondary technical school should give training for entry into industry and commerce at 16+ years to meet demands of local conditions and where possible to include advanced courses for age-groups 16-18 years. It also recommends that these schools should be free from examinations. The subjects of the curriculum are considered individually in Part III.

Technical and Commercial Studies in Secondary Education. Savage, G. London County Council. London, 1950.

The purpose of the pamphlet is to present the substance of the Council's education inspectorate reports on the place of technical and commercial studies in the curricula of London secondary schools at the time of publication.

SUBJECTS OF THE CURRICULUM

(a) General Subjects with a technical or commercial bias.

"*Foreign Languages in a Technical School.*" WALKER, T. H. C. **Voc. Asp.,** Vol. VIII, No. 17. 1956.

A short article surveying the teaching of foreign languages in six local schools. The author includes recommendations for a general policy of teaching languages in the technical environment.

The Teaching of Modern Languages. I.A.A.M. University of London Press. London, 1952.

A general survey of the problems of teaching languages in secondary schools. Various methods of teaching and testing are described and their value assessed. Chapter VII concerns the teaching of modern languages in Technical and Commercial Schools. Because of the state of flux in secondary school organisation, and the need to prepare for Further Education, it is important to devise a good common course for first and second years of the secondary school. The author gives his opinions about what changes are necessary in adapting usual grammar school courses for the purposes of technical schools. Other recommendations include the encouragement of able students to take a second language in the third year, the provision of suitable text-books and the necessity of recruiting teachers who have had experience of vocabularies of technical terms.

"*Standards of Translation.*" NEWMARK, P. **J. of E.,** Vol. 89, No. 1055. June, 1957.

A counsel of perfection is given in matters of translating foreign languages into English. Several points are dealt with thoroughly with examples of translation, including style, phrases, idioms and distinctive grammatical constructions.

"*The Patch Method in Teaching History.*" CARPENTER, P. **J. of E.,** Vol. 88, No. 1047. October, 1956.

This method selects limited periods of history for intensive study by groups of pupils. It is an attempt to solve the second problem stated in the above article, the problem of the present-day practice of condensing history for teaching purposes until it becomes meaningless. The author describes his own method of teaching Elizabethan history to the early years of the secondary school. The disadvantages of this method are also described.

The Teaching of History. I.A.A.M. Cambridge University Press. Cambridge, 1952.

This book covers the whole problem of teaching history. The purpose, methods, visual aids, sources of material, the development of history rooms are fully covered. There are special chapters on sixth form work and on the integration of history with other subjects. No special chapter on technical schools is included but the book is a useful source of information for teacher's of history.

"*Studying the Proper Subject.*" MCNICOL, H. **J. of E.,** Vol. 88, No. 1040. March, 1956.

Often pupils complain of the "uselessness" of learning history. The author assesses various claims for the subject's value. The present methods and syllabus of teaching history are judged and found wanting, particularly in their attempt to teach mainly English history in a continuous process from the Cavemen in the first form to the present century in the last year. The assumptions are attacked that the earlier the historical period the "easier" the teaching content, and that English History is sufficient in itself to teach the mainsprings of human endeavour. A syllabus is given in some detail which recommends that first-year history should be merged with English teaching and should deal with what A. N. Whitehead called the "Age of Romance"; the second year is set aside for a treatment of certain lines of development such as " Man and his Home," power supplies and the conquest of ignorance. The following two years are organised into a course which traces the development of Western Society, its contributions to and its impact upon the family of man.

" *History in the Secondary Technical School.*" Robson, K. J. R., and Briscoe, H. K. **Voc. Asp.,** Vol. VII, No. 15. 1955.

A description of the teaching of History in what was a Junior Technical School at the time of publication of the article. The age-range was 13-15 years. At first Political History was taught but this was replaced by a course which followed the development of themes which were chosen according to the pupils' specialist courses. The syllabus is explained in some detail.

" *Local Sources in the Study and Teaching of Industrial History.*" STONES, E. **Voc. Asp.,** Vol. IX, 1957.

Although the article was written primarily for Technical Colleges, it contains useful hints on where to find sources of local industrial history.

The Teaching of General Science. General Science Sub-Committee of the Science Master's Association. Murray. London, 1950.

This is a general survey of science teaching in secondary schools. Aims, practical work, syllabuses and examinations are dealt with in detail. There is a separate chapter on science for the non-specialist. There is no special chapter about teaching science in technically biased schools, but the book is a useful source of information for all teachers of science.

The Teaching of Science in Secondary Schools. Science Masters' Association and I.A.A.M. Murray. London, 1947.

This includes similar topics to the above book, but has also special chapters devoted to the teaching of science in agricultural areas. Examinations are considered, and there is a section, with examples selected for rural schools, of " new type " examinations.

" *Surveying in the Secondary School Curriculum.*" WATERHOUSE, P. S. **Voc. Asp.,** Vol. VIII, No. 16. 1956.

A justification of the introduction of this subject into the curriculum of secondary technical schools. The reasons for its inclusion are given in some detail.

Handbook for Geography Teachers. General Editor: CONS, J. University of London Institute of Education. Methuen. London, 1955.

In some respects this book and the I.A.A.M. handbook are alternative sources of information. However, each has valuable sections not covered as thoroughly as in the other. In this book are very good book lists for pupils of all forms, and for teachers, classified into the main sections of geographical work. There is a special section on teaching geography in technical schools (pp. 32-37). The writer considers that the syllabus of the first two years should be similar to that given in detail on pages 42-54 in the section dealing with grammar schools. Subsequently, suggestions are made for a specifically technical school course suitable for the age-groups 13-15+ years.

" *Geography in the Education of* 15-18 *Year Olds.*" Memorandum of Geographical Association, presented to the Central Advisory Council published in **Geography**, Vol. 42. 1957.

The Memorandum argues that geography should be included in the curricula of all institutions which provide education for these particular age-groups on the grounds that it is particularly appropriate to the intellectual needs of young people, it presents a balanced view of the world, and is a bridge between the humanities and sciences. Its application to technical school curricula is briefly considered as a practical subject dealing with concrete material which can contribute specialist background courses to commerce, building and engineering, for example.

Local Studies. Prepared under the auspices of Members of the Geographical Association's Standing Committee for Geography in secondary schools. Published by the Geographical Association. Revised edition. London, 1949.

The booklet gives details of the syllabuses and exercises used by a large variety of schools in several districts of the country. Reference is made to some schools which have a distinct technical bias. The maps and diagrams are useful, and there is a good bibliography relevant to the study of local regions in the British Isles.

The Teaching of Geography in Secondary Schools. I.A.A.M. Philips. London, 1954.

This book considers several aspects of teaching geography at the secondary school level. It covers aims, schemes of work, teaching methods, equipment and gives a variety of information about sources of material of great value to all geography teachers. There are special chapters on sixth form work and on geography in secondary modern schools. There is no special section on geography in technical schools, but technical school teachers will still find this book of great value as a source of ideas and information which can be used and adapted for their specific needs.

"*The Teaching of Geography in Schools.*" SCARFE, N. V. **Geography,** Vol. 34, Part 2, No. 164. June, 1949.

This article reviews British research into problems of teaching geography in all types of schools. Although technical schools are not mentioned specifically this summary of research and its recommendations for future inquiries provides a valuable starting point for a programme or original research in teaching this subject in a technical school environment. There is a good bibliography classified under the headings of Teaching Method, Geographical Abilities and Children's Interests and Attitudes.

"*Geology in Schools.*" Report presented by a Research Committee of the British Association for the Advancement of Science at the Annual Meeting held in Sheffield, August, 1956. **The Advancement of Science,** No. 52. March, 1957. Also reprinted separately.

The members of the committee indicate the main advantages of including geology in the syllabuses of both selective and non-selective schools. Although technical schools are not specifically mentioned the application of this subject to this particular type of school is apparent. The article gives no particular schemes of work or of methods of teaching, but its general comments on these topics are valuable. It also discusses the equipment which is necessary at different levels of teaching, and considers the connections with further education in both universities and technical colleges. Careers are also considered. The connections with other subjects such as geography, chemistry, physics and biology are mentioned.

(b) Commercial Subjects

"*A Secondary Course in Commercial Subjects.*" HARTRIDGE, R. J. **Education and Commerce,** No. 19. April, 1957.

The author described a course suitable for G.C.E. (O) level, although this examination is not the prime purpose of the course. It begins in the third form and seeks to use methods which involve the activity of the pupil. It attempts to create links with other subjects and to allow opportunities for the use of typewriters and other office machinery. Visits to factories are encouraged but only after careful planning. The development of the scheme is described in some detail.

Teaching Commercial Subjects—The Modern Approach. INGRAM, B. (ed.). Published by The Faculty of Teachers in Commerce.

Designed for teachers in both secondary schools and Colleges of Further Education, the teaching of commerce, book-keeping, shorthand and typewriting is considered in detail. Bibliographies are included.

Commercial Studies in Secondary Schools. London County Council. London, 1957.

This pamphlet was printed for the consideration of teachers in an attempt to take stock of what is being done, to see how it may be possible to give more purpose to secondary education and encourage higher standards of work having in mind the pupils' needs as adolescents and their future needs as adult citizens and workers. Employment trends in commerce, the importance of general educational attainments, and various commercial subjects of the curriculum are discussed.

Report of Social Committee for Commerce (The Carr-Saunders Report). Ministry of Education. H.M.S.O. London, 1949.

The Report considers the whole question of Education for Commerce, including the secondary school level. Chapter V recommends subjects which it considers to be suitable for inclusion in the curriculum. The Summary and Recommendations are particularly useful.

"*The Teaching of Commerce in Secondary Schools: The Need for Experiment and Discussion.*" WILLIAMS, T. H. **Voc. Asp.** 1948. Vol. I.

The author complains that this problem has been neglected. He analyses six methods of teaching commerce and stresses the need for less passivity on the part of the student and greater integration with other subjects, in attempts to overcome the dislike of most secondary students of this subject.

"*Experiments in the Teaching of Pitman's Shorthand.*" HILL, J. **Voc. Asp.,** Vol. VI, No. 13. 1954.

A short detailed account of an experiment in beginners' shorthand with groups of former secondary modern schoolgirls at a technical college. It attempts to enlist the interest of students by remodelling examples and the content of the course. It is of value for secondary schools.

"*Standardising Shorthand Speed Test.*" JAMES, T. L. (two articles). **Voc. Asp.,** Vols. VIII, IX, Nos. 17, 19. 1957.

These two articles describe a method designed to maintain one desired standard over several years other than by the usual subjective methods commonly used by examiners. The method was developed over sixteen years and has been adopted by New Zealand for Public Examinations in Pitman's Shorthand for speeds of over 80 words per minute.

(c) Arts and Crafts

"*The Cultural Aspect of Machine Drawing.*" BOISSIER, J. **Voc. Asp.,** Vol. VIII, No. 17. 1956.

The author considers that this subject is suitable for any course designed to give a "general liberal education." It analyses the value of the subject to both liberal and vocational courses of study.

Report on the Teaching of Technical Subjects in Secondary Schools. Cmd. 162. H.M.S.O. Edinburgh, 1957. (Published separately from the main Report.)

This section of the main Report briefly considers the contemporary situation of teaching technical subjects in Scottish Junior and Senior secondary schools and makes recommendations for their improvement.

The Trade and Training of the Carpenter and Joiner. H.M.S.O. London, 1933.

Although an older publication, Section III, "The Organisation of Instruction," has comments to make concerning theoretical and workshop instruction and the aims and duration of such courses.

Technical Subjects in Secondary Schools,
Technical Subjects in Secondary Schools (Rooms, Equipment and Safety Precautions).
H.M.S.O. Edinburgh, 1950.

The first pamphlet considers Craftwork, Technical Drawing and Applied Mechanics in some detail as subjects of secondary school curricula. Content, time to be devoted to the subjects and other factors are considered in detail.

The second pamphlet adds supplementary notes on specialist rooms and their equipment.

Handicraft in Senior Boys' Schools (Memorandum VIIIa).
Handicraft in Senior Girls' Schools (Memorandum VIIIb).
London County Council. London, 1933.

For boys the subjects considered are woodwork, metalwork, bookcraft, basketry, leatherwork, weaving, pottery and gardening; for girls, domestic subjects, needlecraft, pottery and gardening. Although these publications are rather old there are some comments and examples which have relevance to present-day conditions.

"*Crafts and the Adolescent.*" Mackley, G. **J. of E.,** Vol. 89, No. 1050. January, 1957.

After a general treatment of the attitudes of adolescents to arts and crafts there is a criticism of arid exercises and a plea for work which "touches school life at many points, brings together many diverse interests and utilizes a wide range of skills." Art, woodwork, needlework, weaving and printing are all briefly mentioned in the development of this theme.

Building Crafts. Min. of Ed. Pamphlet No. 4. H.M.S.O. London, 1945.

A variety of craft courses is considered, including brickwork, masonry, plastering, carpentry, plumbing. The pamphlet is designed for craft courses of Further Education. Some of the sections are of value for those secondary schools which include these vocational subjects in their curricula. For others it will provide information concerning the courses which some of their pupils will take when they leave the secondary school, allowing the provision of background orientation courses at the secondary school level.

Metalwork in Secondary Schools. Min. of Ed. Pamphlet No. 22. H.M.S.O. London, 1952.

The pamphlet discusses the contribution of metalwork to general education as distinct from engineering courses. Reference is made to secondary technical, grammar and modern schools. The emphasis is on general, rather than narrowly vocational education.

" *Woodwork for Girls*." PAULTON, H. E. **Voc. Asp.,** Vol. VII, No. 14. 1955.

The author stresses the need to teach simple woodwork to girls especially, to future housewives. The purpose is essentially utilitarian.

Report on the Teaching of Nautical Subjects in Secondary Schools. Scottish Education Department. Cmd. 162. H.M.S.O. Edinburgh, 1957. (Published separately from the main Report.)

Originally designed for the narrowly vocational training of fishermen, the courses broadened to include nautical subjects in the curriculum of secondary schools where the nature of the environment and community made this desirable. Thus, the vocational aim became secondary. The problems of staffing, accommodation, equipment and subject matter are considered, together with the problem of integrating this with other subjects.

The Teaching of Handicrafts in Secondary Schools. U.N.E.S.C.O. Geneva, 1950. International Bureau of Education: Publication No. 123.

The Report surveys the secondary school teaching of 38 countries, including the United Kingdom. It considers the time devoted to teaching these subjects, the content for both boys and girls, buildings, equipment, etc. Bibliographies are also included.

Education for the Printing and Allied Trades. H.M.S.O. London, 1935.

Several subjects are discussed at length in Part II, including typography, lithography, book-binding, process engraving, electro- and stereotyping. Suggestions for improving contemporary work are also included.

Handbook of Suggestions to Teachers of Printing Subjects. Yorkshire Council for Further Education. Pamphlet No. 54. Leeds, 1955.

Examples are given of suitable courses for " Printing Office Workers," Designers an those working in associated trades and professions. Curricula are considered. It is essent designed for those responsible for younger Part-time Day Release Students in Technical Coll but may also be of value to specialist teachers of the senior age-groups of secondary sch which include these subjects in their curricula.

BUREAU OF EDNL. & PSYL. RESEARCH
LIBRARY

©

National Foundation for Educational Research in England and Wales.
Printed by King, Thorne & Stace Limited, 1-13 Jubilee Street, Brighton, 1,
and published by Newnes Educational Publishing Co., Ltd., Tower House, Southampton Street, Strand, London, W.C.2.,
for The National Foundation for Educational Research in England and Wales, 79 Wimpole Street, London, W.1.